

FACTY FORM 502

N68-15901- N68-15935

(ACCESSION NUMBER)

435

(PAGES)

1

(CODE)

05

(CATEGORY)

NASA CR OR TMX OR AD NUMBER

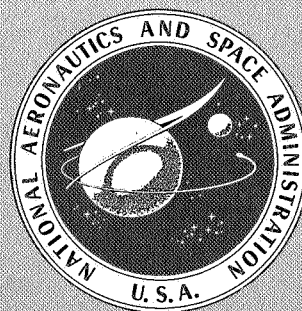
NASA SP-144

THIRD ANNUAL NASA-UNIVERSITY CONFERENCE ON MANUAL CONTROL

UNIVERSITY OF SOUTHERN CALIFORNIA

Los Angeles, California

March 1-3, 1967



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

THIRD ANNUAL
NASA - UNIVERSITY CONFERENCE ON
MANUAL CONTROL

*University of Southern California
Los Angeles, California
March 1-3, 1967*



Scientific and Technical Information Division
OFFICE OF TECHNOLOGY UTILIZATION
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
1967
Washington, D.C.

FOR SALE BY THE CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION
SPRINGFIELD, VIRGINIA 22151 - CFSTI PRICE \$3.00
Library of Congress Catalog Card Number 66-61934

PRECEDING PAGE BLANK NOT FILMED.

Foreword

This volume contains the proceedings of the Third Annual NASA-University Conference on Manual Control held from March 1 to 3, 1967, at the University of Southern California, Los Angeles, California. The program was divided into the following sessions: display, describing function models, decision processes I, computer processing of manual control records, controlled elements, decision processes II, physiological modeling, decision processes III, advanced modeling techniques, and applications. Both formal and informal presentations were made; all of the formal and many of the informal papers are included in this volume.

Contents

	<i>page</i>
Summary -----	1
I. DISPLAY	
1. A Systems Analysis Theory of Manual Control Displays ----- D. T. McRuer and H. R. Jex	9 ✓
2. Two-Dimensional Manual Control Systems with Separate Displays ----- William H. Levison and Jerome I. Elkind	29 ✓
3. Piloted Simulator Display System Evaluation—Effective Resolution and Pilot Performance in the Landing Approach ----- Wendell D. Chase	43 J
4. Holographic Display Systems ----- A. J. Devaney, C. Grauling, and S. Baron	55 ✓
5. Human Operator Describing Functions with Visual and Tactile Displays ----- James C. Bliss	67 ✓
6. A Synthetic Display Technique for Computer-Controlled Simulator and Airborne Displays ----- Jack J. Hatfield	81 ✓
II. DESCRIBING FUNCTION MODELS	
7. Inflight and Ground Simulation Measurements of Pilot Transfer Characteristics in the Compensatory Roll Tracking Task ----- F. D. Newell	99 ✓
8. Describing Functions for Compensatory Tracking of Sine Waves Plus Noise ----- A. C. Beare and A. Kahn	121 ✓
9. A Comparison of Human Response Modeling in the Time and Frequency Domains ----- Lawrence W. Taylor, Jr.	137 ✓
III. DECISION PROCESSES I	
10. Information-Processing Rate as Influenced by the Degree of Response Difficulty: A Discrete Tracking Task ----- Daniel L. Baty	157 ✓
11. A Psychological Approach to Operator Modeling in Manual Control ----- Charles R. Kelley	165 ✓

IV. COMPUTER PROCESSING OF MANUAL CONTROL RECORDS

	<i>page</i>
12. Relationships Between Fourier and Spectral Analyses ----- Lawrence W. Taylor, Jr.	183 ✓

V. CONTROLLED ELEMENTS

13. Manual Control System Performance With Quickened Display, State Variable Display, and Display Gains ----- F. A. Muckler and R. W. Obermayer	189 ✓
14. Sensory-Motor Aspects of Manual Control of High-Inertia Tracking Systems ---- Russell L. Smith and John Lyman	203 ✓
15. Force Feedback Compensation: A New Concept for Improved Manual Control System Performance ----- James H. Herzog and Richard W. Pew	213 ✓
16. Effect of Some Control System Nonlinearities on Single-Loop Compensatory Tracking----- Dunstan Graham	215 ✓

VI. DECISION PROCESSES II

17. An Asynchronous Pulse-Amplitude Pulse-Width Model of the Human Operator ----- M. J. Merritt and G. A. Bekey	225 ✓
18. A Quantal Model for Human Tracking- ----- E. C. Poulton	241 J
19. The Effect of a Random-Sampling Interval on a Sampled-Data Model of the Human Operator ----- G. A. Bekey and J. M. Biddle	247 J

VII. PHYSIOLOGICAL MODELING

20. Peripheral Versus Central Adaptation: Some Preliminary Results ----- G. C. Agarwal, B. Berman, P. Loehnberg, and L. Stark	261 J
21. State of the Art in Developing Models of Skilled Movement ----- C. B. Gibbs	267 J
22. A Neuromuscular Actuation System Model ----- D. T. McRuer, R. E. Magdaleno, and G. P. Moore	281 ✓
23. Pupil Dilation as a Measure of Workload R. O. Anderson and P. E. Pietrzak	305 ✓

VIII. DECISION PROCESSES III

24. "Inhibitory" Control: Concept for a First Model ----- John Lyman and Amos Freedy	311 ✓
---	-------

CONTENTS

	<i>page</i>
25. Supervisory Control of Manipulation -----	315 ✓
T. B. Sheridan and W. R. Ferrell	

IX. ADVANCED MODELING TECHNIQUES

26. Stochastic Modeling of Human Learning Behavior -----	327 ✓
Albert E. Preyss and Jacob L. Meiry	
27. Manual Time-Optimal Control for High-Order Plants -----	351 ✓
Syozo Yasui and Laurence R. Young	
28. Adaptive Finite-State Models of the Human Operator -----	371 ✓
E. S. Angel.	
29. Obtaining Appropriate Human Pilot Describing Functions From Crossover Models and Optimal Control Theory -----	381 ✓
Lee Gregor Hofmann	

X APPLICATIONS

30. Application of Human Transfer Functions to a Design Problem -----	387 ✓
James J. Adams	
31. Asymptotic Stability Studies in Simulated Car Following -----	399 ✓
Robert E. Fenton	
32. Interpretation of Pilot Opinion by Application of Multiloop Models to a VTOL Flight Simulator Task -----	415 ✓
E. W. Vinje and D. P. Miller	
33. Man-Machine System Equalization Determination and its Implementation -----	441 ✓
C. L. Tipton	
34. Applications of the Pilot Transition Response Model to Flight Control System Failure Analysis -----	457 ✓
David H. Weir	

Summary

*George A. Bekey
University of Southern California
and*

*Roger L. Winblade
Office of Advanced Research and Technology, NASA*

The third annual NASA-University Conference on Manual Control was held at the University of Southern California in Los Angeles from March 1 to 3, 1967. Some one hundred specialists from the United States, Canada, England, and Japan attended the conference, which was organized by the authors of this summary.

As contrasted with the previous conference in this series, the major emphasis at the 1967 meeting was on discrete and decision processes in human control actions. A large number of papers were concerned with modeling the decisions which human controllers face in manual control situations. Perhaps this trend will be significant in future research on manually controlled space vehicles. Those tasks where simple, nearly continuous control behavior is possible can be adequately described by means of quasi-linear describing functions. However, much more needs to be learned about situations in which human operators of complex systems are faced with hierarchies of decisions and where their control actions are discrete rather than continuous.

At this conference there was also considerable evidence of interest in displays and in the physiological aspects of manual control. Several papers were concerned with examining the dynamic processes occurring within the visual and neuromuscular systems of the human operator in a manual control task. This, too, may be a trend to be watched in the future.

The formal program was divided into a number of sessions. Highlights of the major sessions are presented in the following paragraphs.

I. DISPLAY

The papers presented in this session covered both studies of conventional display design and an exploration of new techniques. D. T. McRuer et al. presented a systematic approach to the development of displays based on the overall pilot-vehicle characteristics. An interesting study of control systems with separated displays was discussed by W. H. Levison and J. I. Elkind. Their work showed that under these conditions manual control includes both foveal and peripheral vision, the former being used for tracking from one display and the latter for monitoring the second display. Conceivably, this approach could be extended to multivariable display tasks once an adequate scanning model is developed. W. D. Chase reported on a study of the effects of dynamic versus static visual acuity in a fixed-cockpit simulator during visual landing approaches. Pilot-vehicle performance was measured and compared with real flight data. A new concept for implementing a computer controlled electronic display system was presented by J. J. Hatfield. The synthetic display concept utilized is based on the electronic animation of photographs of static-display mockups. The resulting computer driven display is then reproduced in the cockpit by high-resolution television.

Two informal presentations at this session were concerned with novel and unusual display systems. A. J. Devaney et al. discussed the mathematical foundations of holography and

its implications for flight display and flight simulators. This is a challenging area for future displays since they demonstrated that under certain conditions six degrees of viewing freedom can be obtained. J. C. Bliss discussed human-operator describing functions using both visual and tactile displays. The aim of his research has been to obtain a valid comparison between visual and tactile senses in tracking tasks.

II. DESCRIBING FUNCTION MODELS

Quasi-linear describing functions as models for human operators and methods for measuring them have been known for quite some time. A substantial addition to the body of existing data in this field was contributed by F. D. Newell who presented extensive describing function measurements in a compensatory roll tracking task, obtained both in ground simulation and using inflight measurements. A significant conclusion of this work was that it verified the validity of describing function data obtained from ground simulators under stationary conditions. A. C. Beare and A. Kahn presented a criticism of existing describing-function models for the human operator.

Time-domain methods in the analysis of dynamic systems have been utilized increasingly in modern control theory. A particularly interesting feature of the time-domain modeling technique is that it can be extended to the representation of nonlinear behavior on the part of the pilot. An example was presented showing the effect of adding a cubic term to the pilot model on the remnant.

R. L. Stapleford presented an informal discussion concerning pilot describing function measurements in multiloop tasks.

III. DECISION PROCESSES I

In this first session involving human decision processes, three quite distinct papers were presented. D. L. Baty reported on the dependence of the information processing rate of human operators on the degree of task difficulty, in a series of discrete, random input tracking tasks. An index of response difficulty is defined in this paper, and it is shown that the careful design of controllers and switches may have a significant effect on the achievable information-processing rates in pilot-vehicle tasks. C. R. Kelley presented a critique of the describing-function approach to human operator modeling, emphasizing instead a two-time-scale predictive or preview model. Such a model is considered by the author to be more truly representative of actual human-operator behavior. A two-time-scale model is based on a high-speed internal representation or "fast-time model" of the system under control, which can be altered as a result of the observed response, as well as the operator's predictions of future system characteristics, thus taking into account adaptive behavior.

An informal presentation of the human controller as a bang-bang state regulator was given by D. C. Miller. In his experiments the subject was restricted to a three-state controller (+1, 0, -1) for providing inputs to a second-order dynamic system. The subject's task was a minimum-time two-point boundary value problem. The subjects were provided with phase plane displays.

IV. COMPUTER PROCESSING OF MANUAL CONTROL RECORDS

In an evening discussion session three informal presentations were concerned with the use of parameter tracking, hybrid computer methods, and spectral analysis techniques. The

major portion of the evening was devoted to a spirited discussion of the relationships between Fourier analysis and spectral analysis.

V. CONTROLLED ELEMENTS

In spite of its title, the session on controlled elements was in fact concerned with both display and control aspects of manual control. F. A. Muckler presented a comparison of two types of displays on manual control system performance. Both displays included error and error-rate information. One was a conventional quickened display in which the position of a bright horizontal line was dependent on a linear combination of error and error rate. The second, called the "state variable display," presented two bright lines, the positions of which were proportional to error and error rate, respectively. Quadratic performance measures were collected to provide information for synthesis of optimal manual control systems under different performance indices. The paper shows that under certain conditions, the state-variable display could be preferred to the quickened display.

K. Washizu and K. Miyajima in an informal presentation discussed a modification of conventional describing function models for human pilots controlling unstable, second-order systems (second-order systems with negative damping). The authors showed that under these conditions the describing function is modified by the addition of a second-order lead term. The authors hypothesized that this term arises from the pilot's awareness of the existence of periodicity of the response.

Smith and Lyman presented some results of experiments using a high inertia tracking system. The effect of such variables as type of magnification and field of view on integrated absolute tracking error were presented. Two informal presentations were also given at this session. Herzog and Pew discussed the effect of force-feedback compensation on manual control performance. Graham indicated that many common control-system nonlinearities have a negligible effect on human-operator performance.

VI. DECISION PROCESSES II

Several of the papers in sessions VI and VII showed the continuing concern of the manual-control field for mathematical description of discrete decision processes used by human controllers. Merritt and Bekey focused their attention on mathematical models of operator performance in a class of tasks where the operator's response is pulsatile. Such responses are observed commonly in systems of sufficiently higher order. In such systems the response of the human operator may approach that of a bang-bang element. In the class of models proposed here, the human operator was represented as a pulse generator, with each idealized triangular pulse being represented by its time of initiation, width, and amplitude. Regression analysis methods were used to relate these three pulse parameters to the error and error rate in a compensatory task. Likewise, Poulton discussed the need for a discrete or sampling model which would actually mimic human decision processes. He presented only the formal nature of the model, without any quantitative description. A third paper concerned with the sampling behavior of human operators was presented by Biddle et al. The purpose of the study reported in this paper was to examine the effect of random variations in sampling frequency on the power spectra produced by sampled data models of human operators. Careful studies of tracking records have failed to reveal periodicities in the spectra, such as might be expected if human controllers behaved as sampled data systems with periodic sampling. This paper demonstrated that random variation of the sampling interval tends to mask peaks in the output spectra, suggesting that if human operators do sample, they do not do so at a constant rate.

An informal presentation by Smallwood concerned the use of decision theory in the formulation of human-operator behavior, in terms of the encoding of classes of internal models.

VII. PHYSIOLOGICAL MODELING

An interest in a mathematical representation of the "fine structure" of the human operator was evident at this conference. The four papers at this session were concerned with several aspects of biological processes involved in manual control tasks. Three of the papers were concerned with the neuromuscular or output portion of the system and the fourth, an informal presentation, was concerned with the visual or input system.

Agarwal et al. proposed the following hypothesis: that adaptation to stick dynamics takes place via a feedback loop involving the muscle spindle and the spindle afferent nerve, while adaptation to input waveforms takes place via the alpha motor neuron and the alpha efferent nerve, directly to the muscle itself. Thus, it was hypothesized that adaptation to input waveforms is basically an open-loop process in the neuromuscular system. Preliminary results obtained from experiments aimed at isolating the two loops tended to confirm the validity of this hypothesis. Gibbs was likewise concerned with the effects of feedback in controlling movement. He investigated the alternative hypotheses of preprogramed, open-loop, learned patterns of motor control as contrasted with the use of proprioceptive feedback in controlling movement. Specifically, he investigated accuracy of both eye movement and forearm movement in acquisition tasks. The effect of alcohol consumption on the accuracy was also investigated. The major conclusion of his study was that skilled movement requires feedback for its explanation. McRuer et al. presented a detailed neuromuscular system model which was developed in an attempt to provide an analytical explanation of certain characteristics in describing function data, primarily in the so-called "precision model." The paper presented a detailed description of the anatomy of the neuromuscular system, with emphasis on the motor unit and the muscle spindles. A neuromuscular system model including the gamma and alpha motor neuron, the spindle, and muscle dynamics were presented.

An informal presentation by Anderson and Pietrzak indicated that variations in the size of the pupil in the human eye are directly related to tracking difficulty.

VIII. DECISION PROCESSES III

A number of additional topics related to human decision processes were presented at this session. Lyman and Freedy were concerned with a strategy for man-machine interaction, which could be suited for such situations as the control of artificial arms. They suggested a design in which the human operator's adaptive capacity could be used to override preprogramed control movements when necessary. This, they termed "inhibitory control." The remaining papers in this session were informal presentations. Sheridan and Ferrell were concerned with methods of remote manipulation involving both direct manual control and supervisory control involving computer subroutines. Elkind was concerned with the application of signal detection theory to manual control systems, specifically involving transition in controlled element dynamics. Pew et al., in a careful study of low-frequency sinusoidal tracking, suggested that the fundamental characteristic of the neuromuscular system should be regarded as a discrete signal processing system. It was suggested that the motor system generates discrete impulsive changes in force and that error-velocity power spectra reveal two significant peaks which vary as a transport delay is added to the forward control loop.

IX. ADVANCED MODELING TECHNIQUES

The theme of this session on advanced modeling techniques was adaptive and optimal control strategy on the part of human operators. Preyss and Meiry were concerned with a stochastic model of human learning behavior in a manual control task. They selected a task which involved regulating the state of a double-integral plant with a two-position switch. The human operator was conceptualized as a sequential data processing system consisting of a sensor, a decision maker, and an effector. Learning in this model was viewed as an evolution of the probabilities of switching in the state space. Yasui and Young were concerned with manual, time-optimal control of high-order plants. The operators were shown both the instantaneous position-velocity error spot in the phase plane and the projection of the time-optimal switching locus onto the phase plane. The time necessary to reach the equilibrium state for each case was reduced to less than half the time required when no switching locus was displayed. The paper also presented a technique suitable for an analog computer to generate the exact or approximate switching lines. The paper by Angel presented the second phase of the preliminary work discussed at the 1966 conference, in which a finite-state model of a human operator was proposed. The present paper indicated a technique for making this model adaptive by adjustment of threshold levels. This technique made it possible for the model to track ramps of arbitrary slope as well as arbitrary waveforms. An informal paper by Hoffman was concerned with the application of inverse optimal control theory to human pilot adaptation.

X. APPLICATIONS

A series of papers in this session were concerned with the application of manual control theory to problems involving lunar-landing simulators, automobile following, vertical take-off and landing (VTOL) vehicles, and other aircraft. Adams presented the application of describing-function techniques to the analytical design of a manually controlled lunar-landing simulator. This study was particularly valuable since the predicted characteristics of the manually controlled simulator were later verified in practice after the simulator was put in operation. A mathematical model of a display-driver-control stick combination was used by Fenton to study simulated car following situations. In order to obtain the driver-vehicle system which was displaying asymptotic stability under conditions of close spacing between adjacent vehicles, a control stick with a built-in tactile aiding device was developed. Various compensation methods were examined but stability was not achieved. Hence, optimum tracking and asymptotic stability do not occur under identical conditions in the automobile-following problem. Vinje and Miller presented analytical and flight simulator studies which were conducted to develop a mathematical model for interpreting pilot opinion ratings and pilot selection of optimum control sensitivity for a VTOL aircraft hovering task. The task involves control of a coupled two-degree-of-freedom hovering vehicle. It was concluded that quasi-linear pilot models and several analysis techniques are valuable tools for determining pilot-control characteristics and for interpreting behavior. The correlation of pilot-model parameters with pilot-opinion ratings and optimum-control sensitivities was studied.

A number of other applications were discussed in informal papers. Nevins and Johnson discussed the man-computer interface for the Apollo system. Tipton described a technique for man-machine system equalization in order to obtain optimum performance, as specified by the ITAE. The technique was applied to a simulation of a seventh-order system representing manual control of an aircraft rate of descent. Creer and Hardy described a technique for measuring the reliability contribution of a piloted backup system for launch vehicles of the Saturn V class. Weir discussed the application of pilot transition response models to analysis

of flight control system failures. Westbrook described the characteristics of the advanced in-flight simulator (TIFS) now under development, and its potential uses for in-flight measurement of human performance data. Attendees at the conference were encouraged to make use of this forthcoming flying simulation laboratory.

I. DISPLAY

1. A Systems Analysis Theory of Manual Control Displays

*D. T. McRuer and H. R. Jex
Systems Technology, Inc.*

68 15902

This paper summarizes the basic elements of a unified theory for the development of displays for pilots of manually controlled vehicles. The theory combines several recent manual control developments in its structure:

- (1) The vehicle dynamics, environmental disturbances, command structure, and mission criteria are derived in terms of meaningful servo analysis parameters.
- (2) The "best" or "alternative best" feedbacks for the pilot are derived using the "multiloop feedback selection hypothesis," which includes the human operator's describing functions, remnant, and subjective preferences.
- (3) Quantitative evaluation of the system performance measures, information bandwidths, and stability margins are made by systems analysis techniques.
- (4) The required scanning pattern and rates and workload margins are then derived, based on still-tentative pilot monitoring and scanning models.
- (5) The progression and regression of the level of pilot behavior (e.g., during training, transfer, stress, or equipment failure) are treated by the successive organization of perception theory of manual control skill development, and display concepts to enhance the level of display utilization can be quantitatively evaluated.

This theoretical framework provides a paradigm for display development, a rational basis for experimental programs, and a theoretical foundation for analyzing the comparative merits or problems of new operational display systems.

Present-day practice in the development of operational display systems is based largely on intuition and tradition backed by a qualitative understanding of the potentially useful control information. Instrument arrangements or integrated display formats are selected using this experience and background, and are then subjected to exhaustive and expensive development and comparison in manual simulators and, ultimately, are verified in flight. The simulation process usually reveals shortcomings in the preconceived display systems which are overcome by progressive detailed modification and testing. The required modifications can be gross, such as display layout changes or additions to the displayed information; or they can be detailed changes in such things as scaling and dynamic properties. Each progressive modification requires separate evaluation and assessment; and the selection of the best compromise system depends on these assessments. The entire procedure can only be classed as a time-consuming and costly, but currently necessary, cut-and-try process based on tenuous and qualitative guidelines.

The display theory partially exposed here is directed at revolutionizing this art. Its basis is the present analytical understanding of the various ways the pilot can function as a controller and instrument monitor. Putting this understanding actively to work to achieve a detailed, quantitative, analytical delineation of the interactions among the pilot, the display system, and the vehicle allows direct and important savings in the display design/evaluation process.

Central to the evolution of such an integrated display theory is the notion that display configuration design is fundamentally a guidance and control problem which has psychomotor interactions. For example, control analyses using existing pilot models will directly yield the vehicle motion quantities which must be displayed to accomplish some given task, and estimates of the rate at which such displays must be sampled. In turn, the consequences of the eye movements needed for sampling a given display arrangement can be estimated from available engineering psychology data. Expanding on this theme and properly implementing it leads to a practical theory having the following as its primary outputs:

- (1) The quantities or combinations of quantities which must be displayed to enable the specified mission phase or task to be performed. There is seldom a unique, necessary, and sufficient set in complex tasks, so there may be several possible sets of suitable display quantities.
- (2) The pilot's required dynamic behavior in acting on the displayed quantities.
- (3) Performance assessment predictions (e.g., rms system errors and pilot workload data such as information rate, fixation times, and frequencies on each instrument) for each system of displays evolved.

If the theory can be made to deliver these outputs on a reasonably accurate basis, the preliminary design of a display system for vehicular controls would finally become a rational procedure. Experimental studies would still be required to confirm the analytical results, but the systems studied would be far fewer and the experiments more crucial and pointed. Finally, the existence of the theory will make feasible the proper role of analysis and simulation as corroborators and collaborators in the design process.

Following this introduction we shall develop the display theory in two steps, corresponding to the next two sections of the paper. The first describes briefly the components of the display theory. They include the elements of display and their relationship to multiloop pilot control behavior; levels of pilot control, or effective system organization, and their relative merits; the connections between these several forms of pilot/controller characteristics and system organization via the successive organization of perception theory; inducing or retaining high levels of effective system organization; and fixation frequency and duration statistics for the visual subsystem.

The second part combines this potpourri into an over-all theory for display. The basic idea is the following:

- (1) Use closed-loop analysis to predict suitable display quantities and system configurations for various levels of pilot adaptation (i.e., compensatory, pursuit, and precognitive system topologies) and sensory modalities. The result of this procedure is a series of possible competing systems, each capable theoretically of accomplishing the mission tasks.
- (2) Delineate the additional conditions necessary to obtain the assumed level of pilot adaptation for each of the systems set forth above.
- (3) Estimate the properties of the several systems.
- (4) Reiterate as necessary to solve the problem or optimize the design.

It is convenient to describe and, incidentally, to apply the display theory as a process involving two or more stages. In the first-stage version the display theory components are combined for compensatory systems. This is reasonably straightforward to describe, even though some complicating factors are introduced by possible changes in the effective controlled element dynamics using stability augmentation procedures. Still, for a given basic task or mission description and presumed unalterable vehicle dynamics, the first-stage compensatory system provides both minimum requirements and initial points of departure for improvements. The second-stage systems are then evolved as improvement possibilities stemming from the basic compensatory system.

DISPLAY THEORY COMPONENTS

ELEMENTS OF DISPLAY.—In a manual control system the display is the means by which the information required for performance of the task is presented to the operator. Display is here used in the general sense which, in addition to visual stimuli, includes auditory, tactile, and other sensory means of conveying information to the pilot. The display has the following three key elements:

- (1) Forcing function or command-input signals
- (2) Controlled element output signals
- (3) References or background

The forcing function elements introduce exterior information, especially of a nature to show or "command" the desired state of the system. The output signals indicate how the operator's actions are affecting the quantities being controlled (i.e., status information). The background or reference elements are those displayed properties which enable the forcing function command-input signal to be (sometimes) distinguished from the output signal; that is, background is that which enables the behavior of input and output to be separately perceived.

Physically, the three elements of the display may be highly tangible exteroceptive information or, at the other extreme, almost completely interoceptive. Examples of the tangible exteroceptive type are forcing functions generated in avionic systems, background references consisting of wide angle views of the Earth and horizon, and direct visual perception of the aircraft attitude relative to this background. The interoceptively based display elements are exemplified by internally generated command signals or set operating points, proprioceptive indications of control status, and cortico-cerebellar-derived recognition of signal patterns and coherency.

The inputs to the pilot may be perceived by only one of the senses (single modality) or by several senses (multimodality). Fixed-base simulation nearly always *is* limited to the single-modality case (e.g., visual cues only); motion simulators and actual flight involve the multimodality situation (e.g., visual and motion cues). The main justification for ignoring the nonvisual factors in many pilot-vehicle analyses is that the visual cues have been found, empirically, to dominate the pilot's response, and they yield results which seem to be valid. However, there are several situations in which both motion and visual cues are probably important, including pilot-induced oscillations, gust suppression, and spurious motion simulation effects. Multimodality pilot models are in a primitive state of development, and, as yet, few definitive experiments have been performed to show the interactions of the visual display with the motion feedbacks.

The communication effectiveness is maximized by display of only those stimuli necessary for the successful performance of the control task. All too often superfluous or redundant information is provided, tending to saturate the pilot's ability for processing information. Conversely, too little information may unnecessarily burden the pilot with the additional correlations, equalizations, or predictions demanded by the control task. The determination of an optimum balance of information sources and display presentations starts with a knowledge of pilot control theory, that is, the various levels of pilot control action and their relationship to the display.

LEVELS OF PILOT CONTROL ACTIVITY.—In general, there are three essential levels of control activity in manual control systems. These correspond to three different system organizations or topologies which, in turn, delineate the available information paths. A generalized diagram illustrating these possible system organizations is indicated in figure 1. The simplest variant of this system is one which includes only the path indicated by the heaviest line. This elementary compensatory system is a fundamental component of all, or

almost all, man-machine systems. The operator's stimulus consists only of an error signal indicating the difference between the commanded and the actual states of the system. The operator can distinguish between these two states only by intermittent activity; for example, the error equals the system forcing function when the directly controlled output is zero. Fundamentally, the background or reference display element is not present.

Closing (figuratively) either switch 1 or switch 2 changes the system to a pursuit structure. Here, sufficient information is available for the forcing function to be readily distinguishable from the effects of the control responses made by the operator; that is, in essence a background or reference has been added to the display. It is important to note, however, that a pursuit display does not necessarily mean the operator is acting in the pursuit mode. Recent experimental measurements (ref. 1) show that in some cases the operator will act in the compensatory mode even when given a pursuit display.

The third level of pilot activity is the precognitive, or preprogrammed, structure. This can occur in special circumstances in which the predictable aspects of the forcing function are detected and the human plus control system is synchronized to these deterministic forcing-function characteristics. In precognitive control, sufficient predictive aspects are present for the human to develop and then to utilize an entire repertory of preprogrammed responses.

The difference between the three display or system structure organizations described are twofold: The first is available information of a predictive nature; the second is the dynamic performance improvement which can accrue when the information inputs are skillfully utilized. One easily measured dynamic performance difference is an effective time delay between system input and output. For instance, in precognitive circumstances, sinusoidal command inputs can be duplicated at the system output with little or no phase lag and essentially no amplitude ratio decrease for frequencies up to 3 to 4 cps. With random-appearing commands (which preclude precognitive operation), closed-loop pursuit characteristics indicate considerably less effective time delay than present in the closed-loop characteristics of a compensatory system incorporating the same controlled element and operating on the same forcing function. More mundane performance measures, such as mean squared error (or the corresponding rms error, σ_e), are ordered similarly across the three kinds of organization.

MODELS FOR PILOT DYNAMICS.—Analytical models which describe human pilot dynamic characteristics in mathematical terms compatible with flight-control engineering practice exist for each level of control. Not all of these models are based on sufficient experimental data to be definitive. Yet in all cases the models are sufficiently developed to enable the making of relative dynamic performance estimates. A summary is given in reference 2. Here, only the multiloop model will be covered in any detail because the others are already documented elsewhere.

Compensatory-single loop.—By far the most widely and extensively studied pilot characteristics are those for single-loop compensatory systems. Notable early investigations are reported in references 3 to 9. In 1957 these and other data were used to develop a quasi-linear predictive model incorporating a general form and adjustment rules (ref. 9). Since then much additional data have been developed (e.g., refs. 10 and 11), and this model has been substantially extended in accuracy and applicability (refs. 11 and 12).

Compensatory-multiloop.—The evolution of multiloop models for the pilot has proceeded in a way nearly parallel with that for single-loop systems, albeit a few years later. First came studies in which multiloop analyses were made by using hypothetical extensions of the single-loop results (e.g., refs. 13 to 16). Then the results of multiloop experiments (refs. 17 and 18) began to provide a better basis for model building. One of the most interesting and

important findings is that the command-loop describing functions are in excellent agreement with previous single-loop measurements (ref. 17).

The pilot model for multiloop tasks is an extension of the quasi-linear describing function model for single-loop tasks, but with different parameters operating in each loop. Whereas the tradeoffs between performance, stability, and pilot equalization effort are relatively clear-cut in single-loop situations, the number of alternatives becomes much greater in multiloop systems. To converge on the probable feedbacks and equalization selected by the pilot, the adaptive feedback selection hypothesis has been formed, which is based on considerable evidence, and is being validated at the present.

Adaptive Feedback Selection Hypothesis. Given a controlled element having several degrees of freedom, some directly sensed within the general visual field, some observable via visual displays, and (perhaps) some directly sensed using modalities other than vision, then the human pilot evolves, during a learning and skill development phase, a particular multiloop system structure. The active feedback connections in this system are similar to those which would be selected by a skilled controls designer who has available certain variable system characteristics to use for control of given fixed-system characteristics, and who also has available a relative preference guide for the variables. System variables comprise sensing channels for each of the feedback possibilities available to the pilot and possible equalization in each loop which is tailored from an adaptive, but limited, set of equalization forms.

The loops ultimately selected have the following properties :

(1) To the extent possible, the feedback loops selected and equalizer adjustments made will be such as to allow wide latitude and variation in pilot characteristics.

(2) The loop and equalization structure selected will exhibit the highest pilot rating of all practical loop closure possibilities. Preferably, the loops selected can be closed with a pure gain plus a relatively large time delay (on the order of 0.5 sec).

(3) Sampling is minimized; that is, the information used will come from the following sources, in order of preference:

General visual field

Integrated displays

Separately displayed quantities

Adjustment Rules. The outer loops are equalized in accordance with the single-loop rules, with some additions to account for the closure of other (inner or parallel) loops.

(1) Equalization: The equalization form and parameters are based on the "effective controlled element" with the inner loop closed, that is, Y_c includes the augmentorlike action of the pilot in the inner loop. In general, feedback selections which involve much equalization, or different equalization in each loop, seem to be avoided by the pilot. A wide range of acceptable outer loop gains is desirable.

(2) Effective time delay: When several displays must be sampled, an additional increment, $\Delta\tau$ =the effective time delay due to sampling behavior, is added to the basic delay.

(3) Crossover frequency: The same crossover frequency considerations apply as for the single-loop case. Often the outer loop has a lower crossover frequency than the inner loop or single-loop cases because of a larger accumulation of effective time delays.

The inner loops often act as parallel equalization for subsequent loops, or provide feedbacks or crossfeeds which suppress subsidiary degrees of freedom which have undesirable effects on subsequent loops. Because the role of the inner loops is so dependent on outer loop requirements, the rigid rules given above for the outer loop are not generally applicable (e.g., even stability may not be required). The types of inner loops closed and the equalization selected will generally be compatible with one or all of the following considerations:

- (1) Outer loop adjustments per the outer loop adjustment rules become more feasible (e.g., $|Y_p Y_c|$ can be made approximately -20 dB/decade with less outer loop equalization).
- (2) The sensitivity of the closed-loop characteristics to changes in either inner or outer loop pilot characteristics is reduced from that in an outer-loop-only situation. This includes the improvement of gain and phase margins.
- (3) The loop structure and equalization selected are those for which total pilot rating is the best obtainable.
- (4) Time delays due to sampling are minimized.

Insufficient data have been obtained to validate these selection and adjustment rules completely, but one example is available of the pilot's describing function in a simulated multiloop task involving manual control of roll and yaw with ailerons and rudder (ref. 17) which supports the theory.

More recently, data have become available on multiaxis (distinguished from multiloop by the absence of coupling terms in the controlled element dynamics) systems with separated displays (ref. 19), but these have not yet been incorporated into a predictive model.

An essential part of all pilot models is the remnant, that portion of the pilot's output which is not coherent with the input. In single-loop compensatory tasks with conventional controlled elements, the remnant usually comprises a small fraction of the total pilot output and resulting closed-loop errors (ref. 11). In multiloop tasks, there may be an increase in remnant due to residual effects of display scanning, crosstalk, and increased pilot workload. Until definitive measurements are available, one should allow for possibly increased remnant in computation of the closed-loop system characteristics.

Pursuit.—The mechanics of pursuit tracking involve additional pilot operations (over and above compensatory error nulling) on either the input forcing function, the controlled output, or both, depending on the controlled characteristics. For a relatively simple controlled element, pilot operation on the input Y_{p_i} and error functions Y_{p_e} is indicated in the "adjustment rules" of figure 2. In essence, this "implicit pursuit" model assumes that Y_{p_e} is unchanged from the pure compensatory case, and that $Y_{p_i} \doteq Y_c^{-1} e^{j\varphi(\omega)}$ is added to reduce the errors near crossover. Typical experimental data, also shown in the same figure, indicate a good correlation with the model, $Y_{p_i} Y_c \doteq e^{j\varphi(\omega)}$. The source of the phase-angle variations $\varphi(\omega)$ is, as yet, unexplained. Other pursuit data are given in references 1 and 8.

When Compensatory systems are modified in one way or another to pursuit systems, the information available on the system dynamics, performance, and workload must be supplemented by the pilot feedforward dynamics Y_{p_i} in those loops made pursuit. Estimates of average performance, such as the error spectrum and rms level, either can be considered as highly conservative for the pursuit situation, or can be modified somewhat to account for the improved error/input caused by the insertion of Y_{p_i} .

As noted for multiloop situations, there may be an increase in remnant over the compensatory situation as a by-product of the additional equalization required in the adoption of Y_{p_i} , especially since the equalization required for Y_{p_i} and Y_{p_e} is always different. Any increase in remnant due to adopting Y_{p_i} tends to offset some of its performance benefits and may limit the degree of pursuit behavior employed when the compensatory error is already a small fraction of the input (e.g., when $Y_c \doteq K/s$).

Precognitive.—Finally, for conditions in which the pilot can become precognitive, the Y_{p_i} feedforward is effectively all that exists, i.e., $Y_{p_e} = 0$, and the system operates as

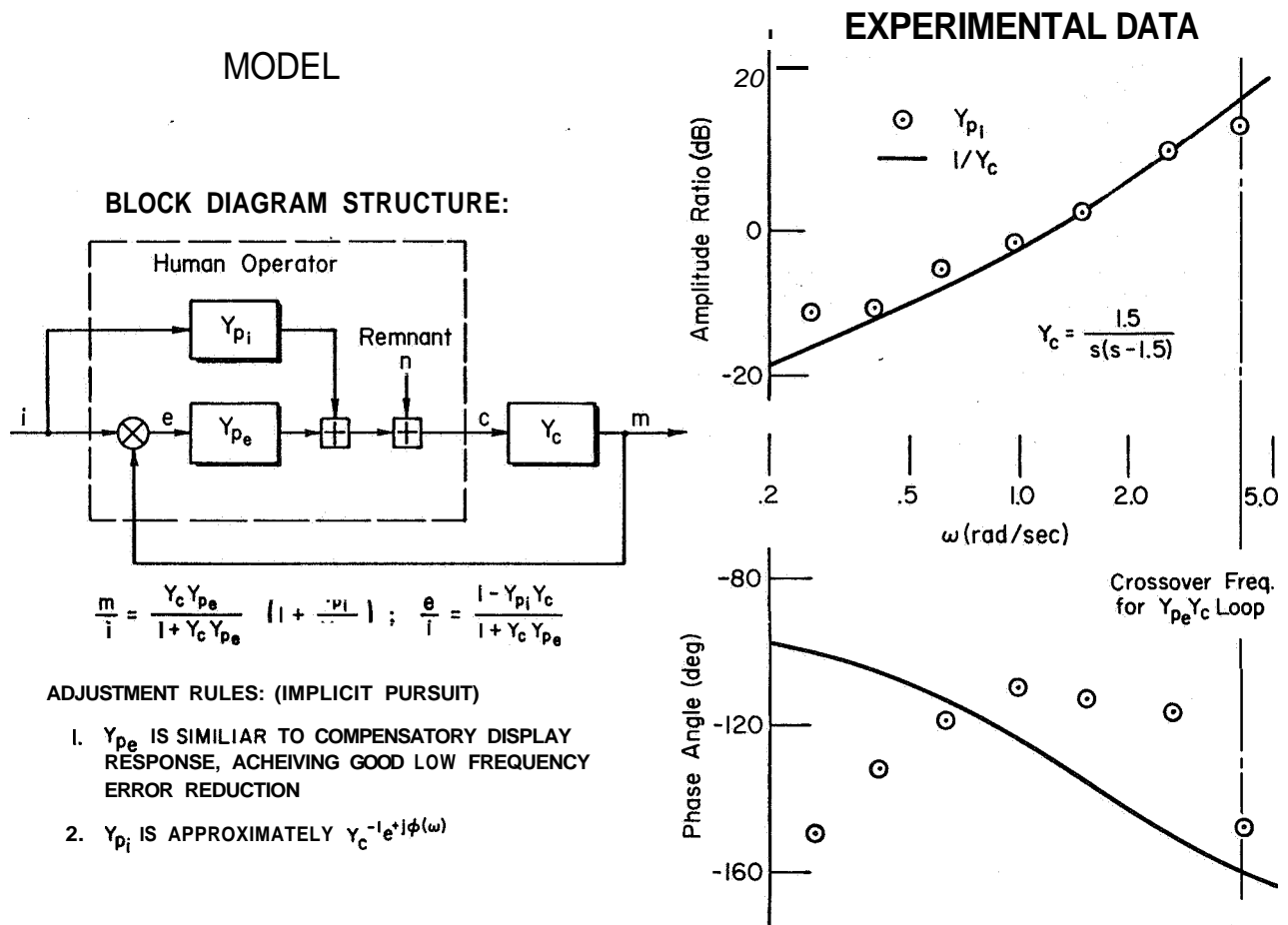


Figure 2.—Example of pursuit measurements.

open-loop insofar as the modality involving the command, i , is concerned. For practical purposes the Y_{pi} adopted is given by $1/Y_c$ so that the transmission from input to output is essentially unity. What actually occurs in the precognitive condition is that the pilot himself generates a patterned or skilled response after just a portion of the completely predictable input has been observed. Thus, what appears on a stimulus-response basis as near-perfect transmission through the system is actually the result of the pilot inserting a programmed command (or series of commands) which, in turn, is triggered by an initial portion of the command-input signal.

Because of the deterministic nature required of the forcing function and the extremely high skill requirements on the operator, precognitive control has many practical limitations. With simple controlled element dynamics such as $Y_c = 1$, precognitive control has been exhibited with sinusoidal, square-wave, and triangular forcing functions, and with more random appearing forcing function waveforms consisting of up to three sinusoids (refs. 20 and 21). There is also some evidence that precognitive operation can be achieved to cope with periodically time-varying controlled elements after much practice (ref. 22).

Derived System Characteristics.—The application of the pilot model for multiloop compensatory systems allows the analyst to estimate, for given forcing function characteristics and a given set of controlled element dynamics, the following:

- (1) Desirable loops to close and, hence, the vehicle motion and system command quantities required for display.
- (2) Pilot dynamics in terms of the pilot describing function for each of the loops closed.
- (3) The closed-loop dynamics such as the system output to command input describing function.
- (4) That portion of the closed-loop average performance due to the pilot's linear operations on the forcing function, that is, σ_e/σ_i . Considerably less accurate estimates can be made for that portion of the system rms error due to the pilot's remnant.
- (5) Closed-loop error spectrum Φ_{ee} including an estimate for the effective bandwidth $\omega_{e\text{eff}}$.
- (6) Ratios of unity-amplitude crossover frequencies for the various loops closed by the pilot which are related to the hierarchy of loop closures and required information rates.
- (7) Relative bandwidths for each of the loops closed, $\omega_{e_j\text{eff}}/\omega_{e_k\text{eff}}$.

In other words, for compensatory display situations at least, a great deal of information about possible system structures, pilot and system dynamics, and average performance can be estimated using a combination of existing theories, validated by experimental data.

SUCCESSIVE ORGANIZATION OF PERCEPTION

The provision of a pursuit-type display is not a guarantee that the pilot will use pursuit control methods. Conversely, a pursuit display is not essential to pilot adoption of effective pursuit tracking. That is, the organization of the system structure is not confined to physical display alone. In fact, one of the most important features of human adaptability is the capacity to set up an internal organization (equivalent to the construction of several signal processing paths within the human) which induces background or references not physically present or which permits highly efficient use of any coherence in the presented stimuli. This capacity to construct a succession of conceptual displays which is effectively equivalent to more elaborate, and often infeasible, physical displays can provide the same type of performance enhancement as that achieved with the physical displays. The above progression in learned use of displays has been termed successive organization of perception (SOP) from reference 23. Similarly, the regression of a system structure from a higher to a lower level of skill (e.g., from an effective pursuit to an effective compensatory condition) is the explanation for many conditions in which system performance is materially degraded.

While it is not possible to evaluate all such SOP implications by strictly analytical means, one can often detect the mode of display utilization being practiced in carefully controlled experiments from the criteria and utilization cited in table 1.

DISPLAY PERCEPTION MODELS

In contrast to the well-established state of the art in overall input-output models for single-loop situations, there are few validated and practical models for the dynamics of the interface between the panorama and instrument displays and the pilot's perception of the relevant signals. Assuming that the desired feedbacks have been obtained from the feedback deletion process described previously, the remaining requirement is to model the scanning frequencies, durations, and patterns.

Ideally, the estimates of pilot stimuli bandwidths obtained using the pilot dynamic characteristics described above can be used to determine estimates for scanning frequencies and durations for the instruments comprising the display system. The connecting links between stimuli bandwidths and scanning frequency durations and patterns are experimental

TABLE 1.—SOP DISPLAY UTILIZATION CRITERIA

SOP Mode	Display utilization	Criteria for detection
Compensatory	Y_{p_e} ; on error only	Y_{p_e} fits well-established laws for compensatory control Input-response delay large
Pursuit	Y_{p_i} ; (mostly) on input Y_{p_e} ; on error Y_{p_m} ; (perhaps) on output	$Y_{p_i} \doteq Y_c^{-1} e^{j\varphi(\omega)}$, with $Y_{p_e} \leq Y_{p_c}$ (compensatory) Lower input-response delays; lower errors (usually) Error display still required for loop stability and residual error control
Precognitive	Y_{p_p} on input cues only; detailed pattern gener- ated within operator	Can remove detailed input, or (for stable Y_c) close eyes, for periods \geq (input- bandwidth) ¹ ; $Y_{p_e} \rightarrow 0$

data on fixation rates and durations as functions of displayed signal properties. Some data pertinent to this problem exist, although they require interpretation before they can be directly applied to manual control. Some of these data will be discussed briefly below. Research in this area is active at several laboratories, so better models are expected in the immediate future.

FIXATION FREQUENCY.—The minimum frequency with which a given instrument must be observed is a basically sampling problem, and the existing sampled data theory is adequate to permit its application to the modeling of experimental data intended for display design and evaluation purposes. A well-known result (Nyquist's sampling theorem) states that a signal with effective bandwidth f_i (in cps) must be sampled at a rate of at least $2f_i$ samples per second to enable the original signal to be reproduced exactly by idealized operations performed on the sampled signal magnitudes only.¹ Since the task of the controller is to reproduce the signal in some sense, then he would be expected to sample the instrument at a rate which is greater than $2f_i$ because his operations cannot be ideal, and the bandwidth cutoff is usually not sharp. In cases where several related instruments require different sampling rates, the minimum scanning rate for the group may be set by the fastest required fixation frequency in the group.

The effect of single-display sampling on operator performance in this context has been studied in a series of complementary experiments accomplished by Vossius and Wezel (refs. 20 and 21). They worked with single sinusoids, harmonic patterns of three sinusoids, and a

¹The general sampling theorem, given magnitude plus R derivatives at each sample, is $\text{bandwidth} = 2f_i / (R+1)$.

random-appearing sum of five sinusoids in single-axis pursuit tracking tasks with $Y_c = 1$. Among their experiments were runs in which the forcing function was initially sampled (presented) at a frequency that was much greater than twice the highest frequency component, followed by a gradual reduction in the sampling rate. The output power spectral density was computed, and at the high sampling rates the operator's output exhibited spectral lines at forcing function frequencies (for forcing function bandwidths within the normal range of manual control system operation). As the sampling rate dropped below two to three times the highest forcing function frequency, the peak at that frequency in the operator's output spectral density disappeared. Similar results were obtained for the other input frequency lines as the sampling rate was reduced further still. Also, new lines appeared in the output spectra which seemed to correspond with the sideband frequencies associated with the sampling operation. These experiments thus clearly demonstrated that the operator was able to reconstruct the forcing function so long as the signal was presented at a rate somewhat greater than $2f_i$ for a signal with bandwidth f_i .

The Vossius-Wezel data were gathered for a single interrupted display. By analogy we can apply these results to multiple displays excited by continuous signals wherein the sampling is provided by the pilot's scanning operations. To the extent that this extrapolation is valid, the fixation frequency on each display that is actually used in control operations must be at least $2f_i$ to $3f_i$. An experimental check of this extrapolation is clearly desirable, bearing in mind the interactions between the minimum required scanning frequencies and convenience of scanning the group.

It is possible that some recent experimental data on interrupted displays in monitoring operations (refs. 24 and 25) can be reinterpreted to provide further fixation frequency data. In these experimental situations, the observer was instructed to push a button whenever any of **six** unrelated instruments displayed a value which exceeded some percentage of full scale. The observer was told that he would be rewarded in accordance with the accuracy to which he was able to detect all of these exceedances. Thus, the task was one of exceedance detection and not signal reproduction or control.

Other factors which will influence the scanning rate are the duration of each sample (if magnitude, rate, and higher derivatives can be detected during each fixation, the signal can be predicted for longer periods between fixations), the urgency of the particular signal, and the demands of other control axes or noncontrol tasks. A grand model to account for all the factors is not yet available.

FIXATION DURATION.—While there are several sets of data (e.g., ref. 25) which lend plausibility to the presumption that fixation duration is a function of permissible error, input amplitude, and other quantities, a fixed duration is a fair first approximation. Consider, for instance, an investigation of standard and advanced panel layouts by Winblade (ref. 26) which resulted in the scan data in table 2 (measured over a 40-sec period). The percentage variation in the conventional panel is at least as great as that evidenced in reference 25 over a variety of conditions. The durations for the advanced panel are fairly constant, however, and have about the same magnitude in time as the mean values in experiment II of reference 25. Thus, variations in duration noted by Senders (ref. 25) are not large in comparison with other sources that are present in display design and evaluation. Thus, the use of a nominal constant duration for all instruments despite signal bandwidth is tentatively indicated, at least until significant relationships with various parameters in the display are more clearly established.

TENTATIVE DISPLAY WORKLOAD CRITERIA.—Using the data and relationships described above, a procedure for evaluation of competing controller-display systems in terms of a "sensory workload" is immediately evident. The bandwidths of the displayed signals are

TABLE 2.—SCAN DATA FOR PANEL LAYOUT

Instrument	Conventional panel			
	Number of samples	Duration, sec	Number of samples	Duration, sec
A	10	0.66	15	0.53
B	3	.63	4	.53
C	2	.50	2	.50
D	7	.45	5	.45

determined by closing the loops with the multiloop quasi-linear operator describing function model. Using the results of Vossius-Wezel, a minimum sampling frequency of twice the bandwidth is applied to each instrument, and the fixation duration is weighted by the type of instrument. The sensory workload W_s can be computed as the total time around the display divided by the total time available (ref. 24), that is,

$$W_s = \frac{\sum d_i}{T_s} = \sum_i 2f_i d_i; \quad T_s \doteq 1/2f_1$$

where f_i is the effective bandwidth of the signal displayed on the i th instrument and d_i is the average fixation duration on that instrument. The potential sensory workload will be less than that actually achieved, for it corresponds to the minimum time required to close the required loops in an ideal allocation of scanning frequency and duration to each instrument. Modifications of the loop closures will change the potential sensory workload and permit the choice of an optimum among competing alternatives. As more data become available, either from new experiments or by reinterpretation of existing material, the tentative procedure described above can be easily modified.

COMBINATION OF DISPLAY COMPONENTS INTO DISPLAY THEORY

COMPENSATORY (FIRST-STAGE) SYSTEMS.—The detailed elements discussed in the last section can be combined to form a comprehensive and quantitative theory of displays used for manual control. The development here will use the block diagram chart of figure 3 and a running description of the steps needed in a typical compensatory application (first-stage analysis).

The information required at initiation of the procedure is a description of the basic task or mission and the environment in which the task is to be performed in systems analysis terms. These provide the command and disturbance structure shown in figure 3, and result in the forcing function specifications. Also needed initially is a detailed description of the vehicle dynamics. For simplicity here, these are presumed unalterable. However, in an actual design problem they can be modified using either force or motion stability-augmenting devices, and with such modification the display requirements can be substantially changed. What is required in practice, therefore, are several iterations using different stability-augmented vehicle dynamics.

Combining the command and vehicle data with the multiloop pilot model results, first, in estimates of the feedback loops closed by the pilot in performing the task. This yields the manual control system topology, and indicates, among other things, the vehicle motion

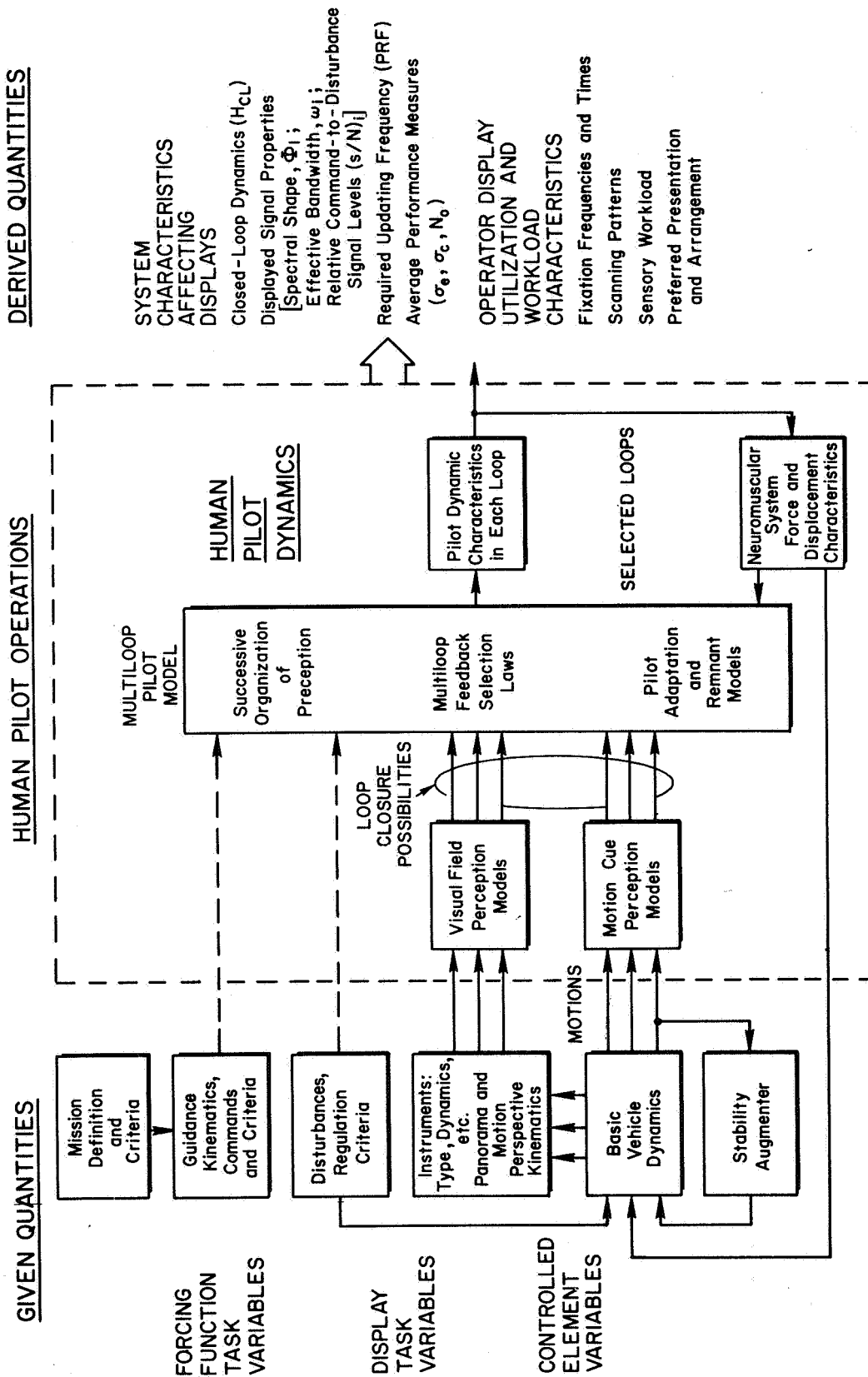


Figure 3.—Elements of a unified manual control display theory.

quantities which must be fed back and displayed for the pilot's use in controlling the vehicle as specified by the command structure. The next steps in the sequence are the determination of the pilot dynamic characteristics in each loop. The dynamics can vary over relatively broad ranges for vehicles with good characteristics, so estimates are made presuming full-attention flight conditions, as they place the greatest demands on the display and pilot. The next step in the procedure is to estimate the closed-loop dynamics using conventional control theory. These, when combined with the forcing function characteristics, provide displayed signal bandwidths and average performance data. Also at this point the minimum required display updating frequencies which are compatible with only minor performance deterioration of the total system can be determined using sampled-data system describing functions in conjunction with those for the pilot. This is a matter of importance when the display system includes digital computing elements, as in the flight-path displays and commands.

Finally, using the sensory workload criterion W_s the optimal display arrangements, fixation frequencies, scan patterns, etc., are determined iteratively using the displayed signal bandwidth estimates and the sampling models mentioned previously. With this step a complete display system has been evolved, and all of its important dynamic entities can be computed.

The first-stage system provides minimum display requirements and will generally exhibit the highest workload and poorest performance. The compensatory system can sometimes be improved by changing the effective controlled element by adopting different feedback paths and/or by the introduction of crossfeeds. Any of these possibilities remain compensatory and can be investigated using fundamentally the same procedure outlined thus far.

SECOND-STAGE IMPROVEMENT POSSIBILITIES.—Any improvement over the best of the compensatory systems requires changes in the level of system organization, that is, by inducing progression along the successive organization of perception scale. Some of the main factors which can help attain and retain higher levels of SOP are as follows:

Provision of a Physical Pursuit Display.—The most obvious improvement is provision of a pursuit system, at least in the command loop. While the system performance is ordinarily improved thereby, this apparently straightforward modification is often physically impossible or otherwise impractical. Some instruments (e.g., an attitude gyro) provide only an error signal, with the command being unavailable as a separate entity. Unwanted inputs, such as gust disturbances, are not measured or displayed. Some of the pros and cons of providing pursuit versus compensatory displays are presented in table 3.

Forcing Function Simplification.—To present the forcing functions in their simplest, most coherent form, the physical forcing functions can sometimes be modified by resolving them into more pertinent axis systems, operating on them with various linear networks (e.g., filtering), and abstracting from the total signal only that information which is most significant.

Preview.—This, in essence, is the same thing as defining system state by its future condition, presuming no nondeterministic changes in current forcing functions. "Preview" can be induced by the use of predictive or fast-time displays and extensive operator training to develop exceptionally high level of task knowledge and skill. The intent of the training is the development of rhythmic, habitual, routine behavior patterns for responses to highly probable inputs. The provision of standardized manipulators and sensing instrument layouts is an important adjunct to advanced training.

Proprioceptive "Display".—As part of the total display, especially designed manipulators can be included to provide force feedback cues as a training aid and/or to provide augmented feedback (one example is a stick-shaker stall indicator). This appears especially promising in high-redundancy operations. For example, pilot operation with a proprioceptive

TABLE 3.—COMPENSATORY VERSUS PURSUIT DISPLAYS

Compensatory (error only)	Pursuit (input and output)
<p><u>Advantages</u></p> <ol style="list-style-type: none"> (1) Simple to build instruments and to simulate; most reliable. (2) Expanded scale possible; easy to detect small errors. (3) Simple to interpret; only one form of action and/or Y_p is required. (4) Can be easily optimized and equalized using pilot-vehicle systems analysis. (5) Only type feasible for pure regulation tasks {e.g., suppress gust disturbances}. <p><u>Disadvantages</u></p> <ol style="list-style-type: none"> (1) Cannot separate disturbances from commands; leads to conservative closed-loop stability criteria and larger tracking errors for low frequency inputs. (2) Error may not agree with secondary cues (e.g., normal accelerations versus \ddot{e}, while tracking a maneuvering target). (3) Command patterns may be masked by remnant-induced errors, thereby impeding improvements due to SOP. 	<p><u>Advantages</u></p> <ol style="list-style-type: none"> (1) Shows more of available information; permits operator to adopt separate control criteria for tracking versus regulation; enhances SOP. (2) Operator can develop a feedforward path (Y_{p_i}) operating directly on the input to minimize the closed-loop errors. Results in reduced τ_{eff} and higher overall system bandwidth, with lower gain and fewer stability problems in the compensatory loop. (3) Improved "conformability," that is, a proper pursuit instrument is a closer analog to the visual field, permitting easier VFR and IFR transitions. <p><u>Disadvantages</u></p> <ol style="list-style-type: none"> (1) Scaling set by largest input command; may result in errors being too small to use effectively. (2) Harder to put inputs on instruments; for example, maneuvering target position in space is hard to derive except from direct visual field. (3) No improvement for regulation against disturbances (for zero input cases). (4) More difficult to interpret and learn; Y_{p_i} and Y_{p_e} involved. (5) More complex to build; less reliable.

display in active parallel control with an automatic landing system would have consequences on training, maintenance of skill, reduction of visual system overloads, and mission reliability.

It is believed that most of the progression from pursuit to precognitive skill levels, for repetitive control actions, is based on the human's ability to execute a triggered pattern of response at the neuromuscular level. Proper control "feel" and force feedbacks can enhance this ability.

Display Content-Minimum Abstraction Levels.—The perspective view as a whole constitutes the input data to the visual system when the real world is "displayed" through a windshield. For instance, the means used to pick out the point toward which a vehicle is moving inherently derives from a two-dimensional abstraction of the outside world as projected on the retina. The probable basis for the perception of visual guidance and control information

is the expansion pattern of the panorama about the "fixed point" toward which the craft is moving. In the parafoveal streamer theory (refs. 27 to 30) direction information is obtained by the observer in motion by making use of the objects in the field of view, which appear to move along paths radiating from the fixed point in the perspective. The tangent vectors to these paths are the "streamers," and the pattern formed by them is the streamer pattern. The instantaneous intersection of their tangents is the point toward which the observer is moving. Sometimes this point can be observed directly, as when the observer's speed is high and objects are close. But usually he has to deduce its approximate position by drawing on the experience accumulated as to the streamer pattern characteristics of various types of motion and various types of ground pattern.

When an artificial display is used, the question arises as to the minimum abstraction required to induce an artificial perspective panorama to provide sufficient streamer cues. Analogs to various minimum "real-world" displays and their performance potential are of very great interest. This problem has yet to be resolved, although some things can be suggested on the basis of existing experience. For instance, in using the real-world display,

(1) Texture is essential to provide a reference frame and a source of streamer points.
 (2) The most important feature of the visual field is the plane of the ground, because it gives the observer the horizon and objects of known size on its surface.

(3) Definite objects are needed in the foreground; without these, as when flying over calm water, flat treeless fields, or featureless sandy desert, the vehicle's height and path cannot be interpreted at all.

(4) Since the retinal image moves as the eye scans the field of view, it is those properties of the retinal image that are invariant to eyeball position which are likely to be useful.

(5) The importance of objects needed to give good motion perspective cannot be determined by static tests of the classical perspective experiments.

Perspective Magnification.—The scaling and magnification of the displayed quantities which constitute the artificial perspective are of profound importance. In landing approach, for example, reference 31 shows that the average touchdown point depends directly on the angular magnification of the perspective. Specifically,

- (1) Compressing the perspective causes steeper approach paths and overshoots.
- (2) Duplicating the perspective results in no touchdown error.
- (3) Expanding the perspective causes shallower approach paths and overshoots.

Framing.—The frame provided by the windshield apparently aids the determination of the relative movement of the spatially summed streamer patterns, thereby improving the pilot's judgment of vehicle motion. Movement thresholds are lower when the size of the visual field is reduced. Questions of how much, tradeoffs, etc., in this connection are not answered, and the phenomenon is not well understood.

Figure-Background Discrimination.—The essence of perception with a display is the emergence of a figure (e.g., an airplane image) from the reference background. Shifting of figure and ground amounts to a sign change, and is a major cause of so-called reversal errors.

Head rigidity and eye fixation, under visual conditions with real-world display, appear to be important factors in retaining a fixed figure-background orientation, and hence in keeping a pursuit or precognitive variety of system behavior.

A wide visual angle also appears to be necessary to avoid a reversal of reference or background (refs. 32 and 33). Such reversals often occur when a pilot switches from VFR to IFR, with its attendant shift in point of view. As long as the pilot can see the earth, he perceives his own aircraft to be banking, climbing, or diving with respect to the stationary earth below. However, as soon as outside vision is excluded, the visible parts of his own craft become the fixed background with reference to which the small moving parts of the instrument

display appear. The pilot sometimes responds in terms of the aircraft as a fixed reference; fortunately he is trained to suppress this innate tendency for control reversal. As another instance, several observers have reported reversals when watching the ground from a banking aircraft. When a large portion of the earth is viewed through a nearby window so that it subtends a relatively large visual angle, the aircraft is perceived to be in a bank; however, when a relatively small area of the earth is seen through a distant window, the aircraft is perceived to be level and the earth tilted. In terms of the SOP structure, these reversals are probably associated with transitions between compensatory and pursuit-like organizations. At transition in a regressive direction, the pursuit loop (from which the stereotyped calibrated repertory of skilled signals derives) will often cause a control movement in the wrong direction.

Display Size and Shape.—Display size and shape effects on inducing or retaining pre-cognition have not been thoroughly examined, although it appears likely that reexamination in this "new" context of old data should provide considerable insight and perhaps factual data. For instance, in the approach situation studied in reference 31, the standard deviation of touchdown error varied directly with field of view. This trend is upheld by both the visual and TV displays used in this series of tests.

System Equalization and Other Existing Techniques.—Use of existing design techniques, such as system equalization and proper use of population stereotypes, are also applicable to the improvement of performance. Quickening is an example of the first, whereas suitable indicator movements typify the second. The typical individual habitually responds to the movements of an indicator by executing a movement in the opposite direction, that is, apparently interpreting the display as if it were an indication of error by responding to it as if to drive the error in the opposite direction. Thus displays in most cases should indicate the direction of the error and not the direction of the movement to be made in correcting it (e.g., the meatball should be high when the approaching pilot is too high).

Superposition of Different Display Fields.—Almost all of the discussion thus far has tacitly presumed that the highly skilled performance desired is developed and retained with one field of view. In the actual circumstances of piloting, however, three fundamentally different fields of view may be present—the instrument panel, the real-world view through the windshield, and the superimposed field occupied by a head-up display. Further, within these there may be narrow and wide subviews, corresponding to the presence or absence of "tunneling" or "fascination" vision. With, then, six possible fields of view, the basic questions of display conformity (among the fields of view) and combination (of information from separate fields) are brought to the fore (refs. 34 to 36). For simple situations field separation is unimportant because simple information can be acquired concurrently from separated fields, for example, by parafoveal vision. For complex situations, and therefore complex fields, concurrent acquisition from both fields can be an efficient process only if the fields are superimposed in the same position. Yet even in this case conformity will not promote combination unless the head-up display exhibits its information in a fashion which can be interpreted by using the same basic rules as are used in the VFR. Pilots are then able to deploy their flying skill in the usual way while drawing information concurrently from both fields.

Some of the factors summarized above are reasonably well understood, and/or extensive ~~and how~~ or systematic data exist to support theoretical developments. Many others, however, are not; it is to some of these that future experimental programs should be addressed.

CONCLUDING REMARKS

The comprehensive theoretical framework shown in figure 3 and described in the foregoing sections provides a paradigm for a quantitative attack on several types of control/display problems. Some examples are as follows:

- (1) Specific display development problems
 - (a) Exposure of potential problem areas and directions for improvement at an early enough stage to minimize detail design risks and costs.
 - (b) Prediction of the best display scaling, filtering, and equalization parameters (e.g., "quickenings").
 - (c) Analysis of display-pilot-vehicle system instability under instrument flight conditions (e.g., flight-path oscillations under ILS guidance).
 - (d) Selection of optimum feedbacks and their gains for integrated displays.
 - (e) Specification of display instrument dynamic range, bandwidth, and tolerable dynamic lags.
 - (f) Estimation of functional limitations on existing instruments as applied to new missions, vehicles, and tasks.
 - (g) Optimization of the location of related instruments for specific mission phases.
 - (h) Rapid post-mortem investigation of anomalies in measured display system performance.
- (2) Rational basis for experiments
 - (a) Analytical evolution of relevant parameters and their interactions, as a guide to the experimental design of display simulations and flight tests.
 - (b) Correlation and unification of the results of numerous ~~ad hoc~~ display tests performed in the past.
 - (c) Exposure and prediction of solutions to new problem areas in integrated displays, terrain-following displays, etc.
 - (d) Evolution of display simulator functional requirements (e.g., instrument servo lags, and CRT capabilities).
 - (e) Interpretation of experimental findings in an analytical manner to permit their extrapolation to actual flight and future problems.
- (3) Evaluation of new or operational display systems
 - (a) Assessment of competing display systems by a common set of ground rules, models, and assumptions (e.g., via a standard set of forcing functions, pilot models and adjustment rules, and sensory workload criteria).
 - (b) As a guide for developing comparative tests in simulators or in flight.
 - (c) Reduction of different evaluation procedures to a common basis.
- (4) Generation of new display concepts and improvements
 - (a) Establishment of a rational basis for minimal elements on contact-analog displays.
 - (b) Evolution of displays to resist disorientation and to achieve optimum head-up display arrangements.
 - (c) Synthesis of improved blind-landing and terrain-following displays.
 - (d) Recommendation of the best training and utilization procedures to enhance learning and skill transfer.

The foregoing list is not intended to suggest that the proposed theory is a panacea for all display problems, but it does offer the promise that a great number of display problems which are presently solved by expensive cut-and-try procedures can be handled more rationally and efficiently in the near future.

REFERENCES

1. Wasicko, R. J.; McRuer, D. T.; and Magdaleno, R. E.: Human Pilot Dynamic Response in Single-Loop Systems with Compensatory and Pursuit Displays. AFFDL-TR-66-137, Dec. 1966.

2. McRuer, D. T.; and Jex, H. R.: Effects of Task Variables on Pilot Models for Manually Controlled Vehicles. Presented at the AGARD Specialists' Meeting on Stability and Control (Cambridge, England), Sept. 26, 1966.
3. Tustin, A.: The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design. J. IEEE, vol. 94, part IIA, no. 2, 1947.
4. Russell, Lindsay: Characteristics of the Human as a Linear Servo-Element. Master's Thesis, MIT, Dept. of Elec. Eng., May 18, 1951.
5. Anon.: Human Dynamic Study. Rep. GER-4750, Goodyear Aircraft Corp., Apr. 8, 1952.
6. Anon.: Investigation of Control "Feel" Effects on the Dynamics of a Piloted Aircraft System. Rep. GER-6726, Goodyear Aircraft Corp., Apr. 25, 1955.
7. Krendel, E. S.; and Barnes, G. H.: Interim Report on Human Frequency Response Studies. WADC-TR-54-370, June 1954.
8. Elkind, J. I.: Characteristics of Simple Manual Control Systems. TR-116 MIT Lincoln Lab., Apr. 6, 1956.
9. McRuer, Duane T.; and Krendel, Ezra S.: Dynamic Response of Human Operators. WADC-TR-56-524, Oct. 1957.
10. Hall, I. A. M.: Effects of Controlled Element on the Human Pilot. WADC-TR-57-509, Aug. 1958.
11. McRuer, Duane; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems—Theory, Models, and Experiments with Controlled Element and Forcing Function Variations. AFFDL-TR-65-15, July 1965.
12. McRuer, Duane T.; Graham, Dunstan; and Krendel, Ezra S.: Manual Control of Single-Loop Systems: Part I. J. Franklin Institute, vol. 283, no. 1, Jan. 1967, pp. 1-29.
13. Cromwell, C. H.; and Ashkenas, I. L.: A Systems Analysis of Longitudinal Piloted Control in Carrier Approach. Tech. Rep. 124-1, Systems Technology, Inc., June 1962.
14. Ashkenas, I. L.; and Durand, T. S.: Simulator and Analytical Studies of Fundamental Longitudinal Control Problems in Carrier Approach. Proc. AIAA Simulation for Aerospace Flight Conference, Aug. 1963.
15. Durand, T. S.; and Ashkenas, I. L.: Theory and Simulation of Piloted Longitudinal Control in Carrier Approach. Rep. 130-1, Systems Technology, Inc., Sept. 1963.
16. Stapleford, Robert L.; Johnston, Donald E.; Teper, Gary L.; and Weir, David H.: Development of Satisfactory Lateral-Directional Handling Qualities in the Landing Approach. NASA CR-239, 1965.
17. Stapleford, R. L.; McRuer, D. T.; and Magdaleno, R.: Pilot Describing Function Measurements in a Multiloop Task. NASA CR-542, 1966.
18. Todosiev, E. P.; Rose, R. E.; Bekey, G. A.; and Williams, H. L.: Human Tracking Performance in Uncoupled and Coupled Two-Axis Systems. NASA CR-532, 1966.
19. Elkind, Jerome I.; and Levison, William H.: Results from Studies of Two-Variable Manual Control Systems. Presented to AGARD Ad Hoc Panel on Guidance and Control Symposium on The Human Operator and Aircraft and Missile Control (Paris, France), Sept. 5-6, 1966.
20. Wezel, F.: Untersuchungen über die Willkurbewegung der Menschlichen Hand Mit Getastet Dargebotenen Reizmustern. Thesis, J. W. Goethe Univ., Frankfurt am Main, 1962.
21. Vossius, G.: Die Vorhersageeigenschaften des Systems der Willkurbewegung. Nenre Ergebnisse der Kybernetik, R. Oldenbourg, Munich, 1964.
22. Hess, R. A.: The Human Operator as an Element in a Control System with Time-Varying Dynamics. AFFDL-FDCC-TM-65-34, June 1965.
23. Krendel, Ezra S.; and McRuer, Duane T.: A Servomechanisms Approach to Skill Development. J. Franklin Institute, vol. 269, no. 1, Jan. 1960, pp. 24-42.

24. Senders, John W.: The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems. IEEE Trans. on Human Factors in Electronics, vol. HFE-5, no. 1, Sept. 1964, pp. 2-4.
25. Senders, J. W.; Elkind, J. I.; Grignetti, M. C.; and Smallwood, R. D.: An Investigation of the Visual Sampling Behavior of Human Observers. Rep. 1246, Bolt Beranek and Newman, Inc., May 10, 1965.
26. Winblade, Roger L.: Current Research on Advanced Cockpit Display Systems. AGARD rep. no. 491, Oct. 1964.
27. Gibson, James J.: The Perception of the Visual World, Houghton Mifflin Co., 1950.
28. Calvert, E. S.: Visual Judgments in Motion. J. Institute of Navigation, vol. VII, no. 3, July 1954, pp. 233-251.
29. Havron, M. Dean: Information Available from Natural Cues During Final Approach and Landing. Rep. no. HSR-RR-62/3-MK-X, Human Sciences Research, Inc., Mar. 1962.
30. Gordon, Donald A.: Perceptual Basis of Vehicular Guidance. Bureau of Public Roads, vol. 34, no. 3, Aug. 1966, pp. 53-68.
31. Kibort, Barnard R.; and Drinkwater, Fred J., III: A Flight Study of Manual Blind Landing Performance Using Closed Circuit Television Displays. NASA TN D-2252, May 1964.
32. Moore, A. D.: Perceptual Disorientation During Landing of Airplane. Science, vol. 92, no. 2395, Nov. 22, 1940, pp. 447-478.
33. MacLeod, R. B.: Comments on Article by A. D. Moore, Spatial Disorientation During Landing of Airplane. Science, vol. 92, no. 2400, Dec. 27, 1940, p. 607.
34. Naish, J. M.: Combination of Information in Superimposed Visual Fields. Nature, vol. 202, no. 4933, May 1964, pp. 641-646.
35. Perry, D. H.; and Naish, J. M.: Flight Simulation for Research. J. Roy. Aero. Soc., vol. 68, no. 646, Oct. 1964, pp. 645-662.
36. Naish, J. M.: Display Research and Its Application to Civil Aircraft. J. Roy. Aero. Soc., vol. 69, no. 658, Oct. 1965, pp. 662-679.

2. Two-Dimensional Manual Control Systems with Separate Displays

*William H. Levison and Jerome I. Elkind
Bolt Beranek and Newman, Inc.*

768 15903

The results of a current study of multi-variable manual control systems are presented. The objectives of this study are to investigate the human controller's behavior in multivariable control situations and to develop models of the controller which take into account both the monitoring and the control functions that he typically performs in such systems.

A series of two-variable manual tracking experiments was performed in which subjects were required to view two separated displays and operate two control devices to control the system. Performance was measured as a function of the display separation, the forcing function bandwidth, the task difficulty, and the controlled-element dynamics. Human controller describing functions, eye-movement distributions, and normalized mean-squared tracking error were obtained. Measurements of human controller describing functions when a single display is viewed peripherally for control of a single variable system were also obtained.

A model for the human controller in the two-axis control situation was developed. It was tested against the data and found to be a good predictor of performance. Extensions of this model to higher dimension systems are discussed.

In this paper we present and discuss some of the results we have obtained in a current experimental and theoretical study of multivariable manual control systems. This study has two principal objectives. The first is to determine the extent to which existing models (refs. 1 and 2) for the human controller, which have been developed from studies of single-variable manual control systems, predict human controller behavior in more complex multivariable systems. The second, and more important, objective is to develop new human controller models which take into consideration both the monitoring (ref. 3) and the control functions performed by the human operator in most real systems and which will predict with good accuracy human controller behavior in realistically complex systems. We have begun our research program with studies of two-variable systems because they contain two essential features of more complex control systems—multiple axes of control and sharing of visual attention—and yet are simple enough to permit detailed and carefully controlled experimental study.

Most of the experiments reported in this paper were structured so that comparisons could easily be made of the controller's performance on the one-axis foveal task (the conventional single-variable tracking situation), the one-axis peripheral task, and the two-axis task. Primary experimental variables were (1) the display separation, (2) the forcing-function cut-off frequency, (3) the differences between the left- and right-hand tasks, and (4) the controlled-element dynamics.

Most studies of multi-axis manual control systems have focused on measures of system performance, such as mean squared tracking error, rather than on descriptive measures of the human controller (refs. 4 to 7). Two-axis human controller describing functions have

been reported only recently, and these were obtained in control situations requiring no sharing of visual attention (refs. 8 to 10).

A comprehensive study of the human controller of a two-axis system with integrated controls and displays has been previously reported by the authors of this paper (Levison and Elkind, ref. 10). In that study we investigated the effects of input bandwidth and of task differences on the relation between one-axis and two-axis tracking performance. The task differences investigated were (1) different input bandwidths on the two axes, and (2) different controlled elements on the two axes.

We found that after considerable training the subjects were able to track two axes almost as well as one axis when the control situation was homogeneous. The two-axis normalized mean squared tracking error (NMSE) was on the average about only 10 percent greater in the two-axis situation, and the one- and two-axis describing functions were nearly identical. Bandwidth had no consistent effect on these relationships. Most of the increase in error could be attributed to an increase in remnant which we think may have resulted from an inadvertent and random coupling of the movements between the two axes. Since the source of this coupling could have resided in the visual system, the motor system, or in the intervening central pathways, we called this effect "visual-motor interaction."

When the plant dynamics were K on one axis and K/s^2 on the other, large increases in NMSE and appreciable changes in describing functions were observed. (These observations are consistent with the results of others (refs. 5 to 7).) In addition, we observed that the describing function on the K axis appeared to adopt some of the characteristics of the describing function on the K/s^2 axis. The requirement of the human controller to generate simultaneously two kinds of equalizer Characteristics thus appeared to be a direct source of performance degradation.

Wierwille and Gagne (ref. 11) measured time-varying transfer characteristics (mathematically equivalent to describing functions) of the human controller in one- and two-axis tracking situations. The two-axis tracking error was displayed oscillographically in one experiment via a single dot free to move in two dimensions, and in another by 2-meter movements separated by an amount sufficient to require visual scanning. The authors found that the time-variability of the transfer characteristics was greater when the displays were separated than when they were integrated.

Fitts and Simon (ref. 4) also investigated a two-axis tracking situation in which the tracking errors were provided by two spatially separated meter movements. Time-on-target performance was continuously degraded as the display separation increased. (No descriptors of the human controller were computed.) The authors concluded that peripheral vision provided information for accurate eye movements and possibly for control movements.

Other investigators have shown the importance of peripheral vision in purely monitoring tasks. Sanders (ref. 12) and Senders (ref. 13) have both found that (1) signal detection is better than chance for signals located as much as 80° into the periphery, and (2) reliability of detection decreases monotonically as the peripheral viewing angle is increased. Furthermore, the ability to utilize information from a signal source located at a given angle into the periphery depends on the precise location of the source and upon the characteristics of the signal (refs. 14 and 15).

Sanders has proposed the following three functional levels of the visual field: (1) the stationary field, the display angle within which a task can be performed via peripheral vision; (2) the eyefield, in which eye movements (but not head movements) are necessary for satisfactory performance; and (3) the headfield, in which head movements are required.

From these studies of multivariable control and monitoring situations we may reasonably predict the following characteristics of two-axis manual control systems: (1) two-axis performance will be essentially the same as one-axis performance if the control situation is

homogeneous and if the controls and displays are integrated; (2) two-axis performance will be degraded if the dynamics on the two axes are different; (3) two-axis performance will be degraded if the displays are separated so that only one can be viewed foveally; (4) information obtained peripherally will be important when the displays are separated; and (5) the time-variability of the controller's describing function may increase when the displays are separated.

EXPERIMENTAL CONTROL SITUATION

In a typical multivariable manual-control situation several displays are used to present information to the human controller about the state of the system, several state variables have to be controlled, and several control inputs must be provided by the human controller by means of one or more control devices. We would expect, *a priori*, that the performance of the human controller in this kind of situation would be affected strongly by factors such as the kind, number, and location of the displays, the spectral composition of the command and disturbance inputs to the system, the number of control inputs that must be produced by the controller, the dynamics of the system being controlled, the performance constraints that must be satisfied, and the cost function that is minimized. Our two-variable experiments were designed in a manner that allowed us to investigate many of these factors.

The apparatus we used was of conventional design. Information was presented in a compensatory format via cathode ray tubes. Two error dots were displayed, one for each axis, and both dots moved only vertically. A single cathode ray tube was used to present both error displays when a very small separation between displays was desired. Two cathode ray tubes were used to provide display separations of 30' and 56'.

The hand controls were modified versions of the Measurements Systems, Inc., Model 435 force-sensitive control. A flexible shaft was attached to the control to provide the subjects with both force and displacement feedback. The control was essentially spring restrained, omnidirectional, and had no perceivable backlash or friction. Two controls were used, one operated by the subject's left hand and the other by his right hand. The control display relationship was compatible: vertical movements of the left control affected the vertical motion of the error dot on the left display and the right control affected the right display.

The input disturbances were pseudo-Gaussian noise generated by summing 17 sinusoids whose spacings and amplitudes were adjusted to approximate a rectangular spectrum plus a high-frequency, low-amplitude shelf. The input bandwidths were between 0.5, 1.0, and 2.0 radians/sec. The inputs in the two axes were uncorrelated. The controlled-element dynamics in the two axes were uncoupled and were of the forms K , K/s , and K/s^2 .

The input, error, control, and output signals were recorded on magnetic tape in analog and digital format. Subject eye movements were obtained by sensing the electro-ocular potentials and were analyzed to determine the distribution of observation times. Human controller describing functions were computed using techniques described elsewhere (refs. 16 and 17). In addition, mean-squared errors and mean-squared remnants were computed.

The subjects were instructed to minimize the mean-squared tracking error when tracking a single axis. When tracking two axes, the subjects were instructed to minimize the sum of the mean-squared errors. The subjects were trained to an apparently asymptotic performance level on each of the conditions investigated.

EXPERIMENTAL RESULTS

EFFECTS OF DISPLAY SEPARATION.—Peripheral vision is an important factor in the control of the system when the displays are separated by more than a few degrees (ref. 4). Accordingly, one of the first experiments we performed was to investigate the human controller's dynamic response characteristics when the display was viewed peripherally. In these experiments, the subject was asked to fixate on one display whose error dot was stationary and to track the error dot that was displayed in the peripheral region of his visual field. The separation between these displays was 30° and 56° of visual angle—a range that spanned most of the "eyefield" (ref. 12). In addition, the controller's two-axis performance was obtained for display separations of 0.8° , 30° , and 56° , and his one-axis performance was measured when the error dot was viewed foveally. The input disturbance had the same bandwidth in both axes, 2.0 rad/sec, and both controlled-element dynamics were K/s.

Average human controller describing functions and remnant measurements for the one-axis and two-axis tasks are shown in figure 1. The describing functions have been approximated by transfer functions of the form

$$H(j\omega) = K_h \frac{\left(1 + j \frac{\omega}{\omega_2}\right) e^{-j\omega\tau}}{\left(1 + j \frac{\omega}{\omega_1}\right)} \quad (1)$$

where K_h is the low-frequency gain, τ is the controller's effective time delay, and ω_1 and ω_2 are parameters of the equalizer that the controller apparently used to optimize performance (ref. 2). The parameters of the transfer functions are tabulated in table 1.

TABLE 1.—EFFECT OF DISPLAY SEPARATION ON THE ANALYTIC APPROXIMATIONS TO THE HUMAN CONTROLLER'S DESCRIBING FUNCTION

Separation, deg	Mode of tracking	K_h , dB	ω_1 , rad/sec	ω_2 , rad/sec	τ , sec
0	1-axis foveal	2	4	10	0.12
30	1-axis peripheral	-5	-	-	.20
56	1-axis peripheral	-9	-	-	.30
0.8	2-axis	-1	4	8	.15
30	2-axis	-1	4	8	.15
56	2-axis	-6	-	-	.20

The most important effect of a change in display separation on the describing function, from the point of view of system performance, was a decrease in K_h with increasing separation. As the separation was increased from 0° to 56° , the one-axis and two-axis K_h decreased by about 11 dB and 5 dB, respectively. The one-axis effective time delay increased from 0.12 to 0.3 second, whereas the two-axis delay varied only from 0.15 to 0.2 second. The mid-frequency lag-lead behavior evident in both the one-axis and two-axis describing functions for small separations disappeared with increasing separation.

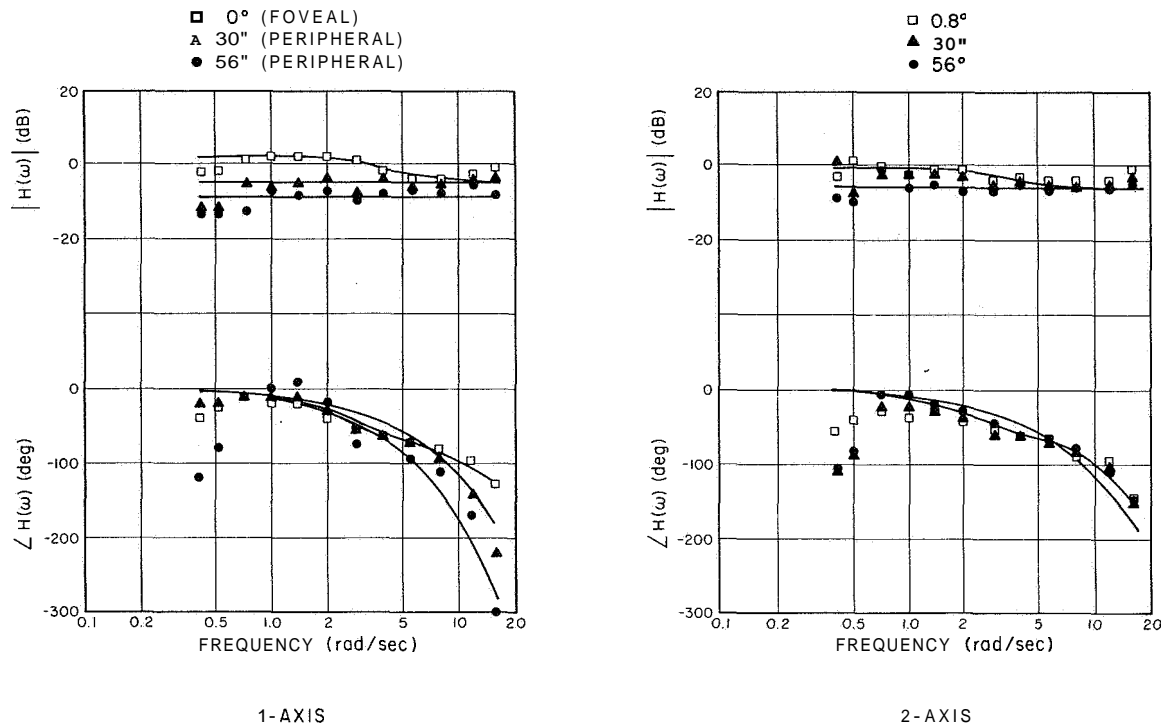


Figure 1.—Effects of display separation on average human controller describing functions for four subjects. Controlled element dynamics were K/s on both axes, and input bandwidths were 2 rad/sec on both axes. 0 dB = 10^5 dynes/cm.

Table 2 shows the fractional remnant powers for both the controller's output (the "stick") and the tracking error. Remnant is defined as the fraction of signal power that is not linearly correlated with the forcing function.

Separation, deg	Remnant stick			Remnant error		
	1-axis foveal	1-axis peripheral	2-axis	1-axis foveal	1-axis peripheral	2-axis
0.8	0.17	--	0.19	0.33	--	0.34
30	.17	0.48	.40	.33	0.70	.53
56	.17	.65	.54	.33	.67	.58

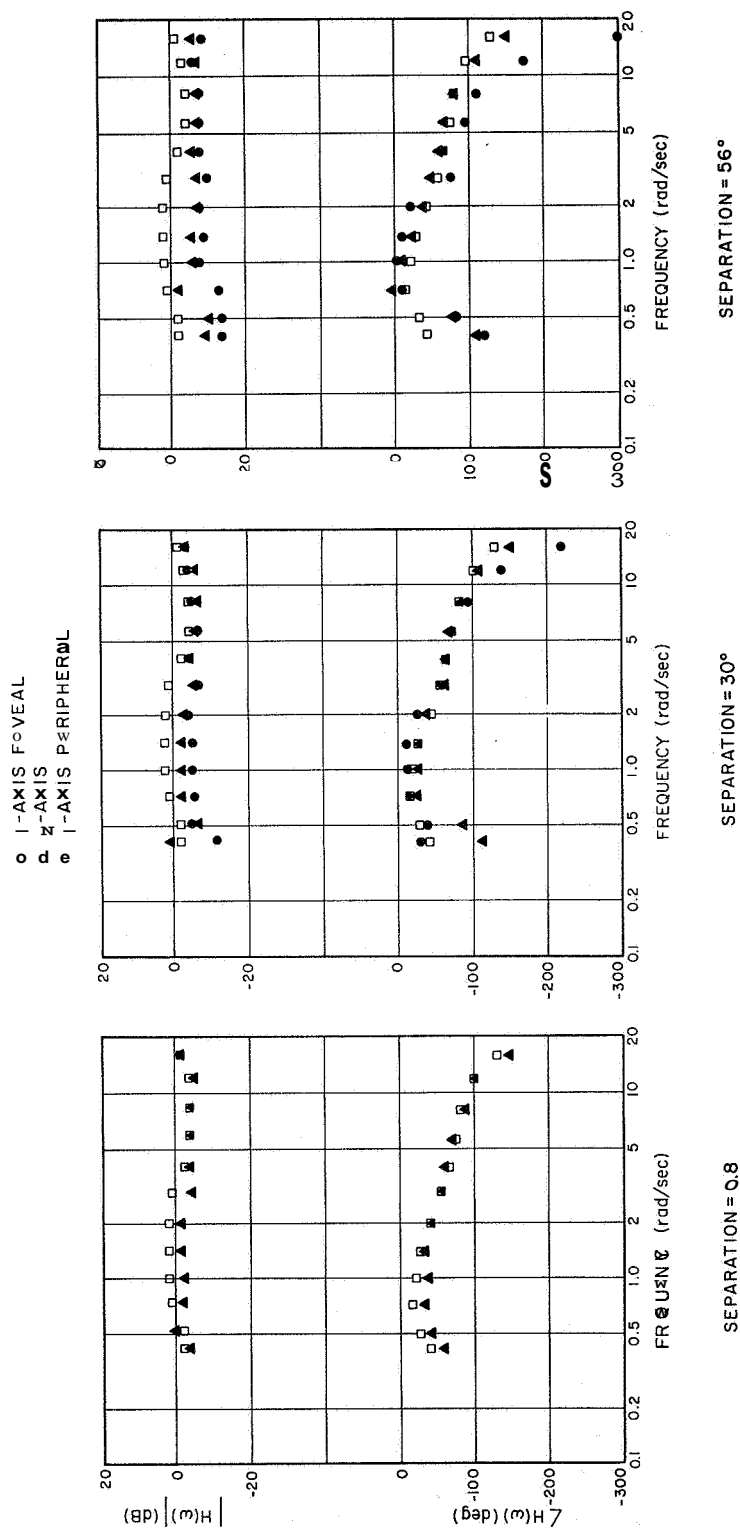


Figure 2.—Comparison of one-axis foveal, one-axis peripheral, and two-axis average human controller describing functions with separated displays for four subjects. Controlled element dynamics were K/s on both axes, and input bandwidths were 2 rad/sec on both axes. 0 dB = 10^5 dynes/cm.

Note that the shape of the describing function was essentially invariant with respect to the mode of tracking—especially for the 0.8° and $30'$ separations—and that the two-axis amplitude ratio and phase shift curves fell between the corresponding one-axis foveal and one-axis peripheral curves.

The mean squared error scores for the one-axis peripheral task (MSE_p) and the two-axis task (MSE_2), normalized with respect to the scores for the one-axis foveal task (MSE_f), are given in table 3. The mean-squared error behavior was consistent with the describing function behavior—the error increased as the display separation increased. For separations of $30'$ and 56° the two-axis MSE was about equal to half the peripheral MSE.

Display separation, deg	MSE ratios		Mean observation time, sec
	MSE_2/MSE_f	MSE_p/MSE_f	
0.8	1.5	--	--
30	3.9	7.5	1.2
56	8.9	17	1.3

Since both error dots could be viewed essentially foveally when the separation was 0.8° , the MSE_2/MSE_f ratio of 1.5 observed with this separation indicates the performance degradation caused solely by the requirement to track two axes simultaneously. This increase in error was larger than was observed with the integrated display experiments reported previously (ref. 10). The greater increase between one- and two-axis tracking in the separated display case is consistent with results reported in the literature (ref. 5), and may result from the fact that coordinated movements cannot be made when the displays are separated. The MSE_2/MSE_f ratios of 3.9 and 8.9 observed for the larger separations show additional performance degradation due to visual scanning.

In spite of the fact that the human controller's describing function has very low gain during peripheral viewing, peripheral vision plays an important part in control. The peripheral gain is still high enough to permit useful control activity. We were able to demonstrate the importance of peripheral viewing in some informal experiments in which we blanked the display that the subject was not looking at. We found that under these conditions the subjects were not able to keep the error dots on the displays for the duration of the tracking run.

Mean observation times were computed from recordings of eye movements that were obtained when the display separations were $30'$ and $56'$. The observation time is the time interval that begins when the subject moves his eyes from one display to look at the other and ends when he moves them away from the second display. The observation time therefore includes both a movement time and a fixation time. Eye movements were not required when the separation was 0.8° , and we have no information about the visual sampling behavior of the controller in this situation.

Mean observation times are shown in table 3. The mean observation time increased from 1.2 to 1.3 seconds as the display separation was increased from $30'$ to $56'$. This increase may represent the additional time required for the eyes to travel between the displays. The observation times obtained in this experiment are about three to four times longer than those observed by Senders (ref. 3) in a pure monitoring task.

EFFECTS OF INPUT BANDWIDTH.—The input bandwidth was the experimental variable in the second experiment. Bandwidths of 0.5, 1.0, and 2.0 rad/sec were used. The bandwidth was the same on both axes, and the controlled-element dynamics were K/s. The display separation for this experiment, and for all further experiments, was 30°.

Mean-squared error ratios and mean observation times are given in table 4. The MSE_2/MSE_f and MSE_p/MSE_f ratios increased between 10 and 20 percent as the bandwidth was increased from 0.5 to 1 rad/sec and increased by factors of 1.5 and 1.9; respectively, as the bandwidth was further increased to 2 rad/sec. Corresponding decreases occurred in low-frequency gain. Except for slight gain differences, the relation between the one-axis foveal, one-axis peripheral, and two-axis describing functions was essentially invariant with bandwidth and was the same that we observed in the preceding experiment. Bandwidth had no consistent effect on remnant.

Bandwidth, rad/sec	MSE_2/MSE_f	MSE_p/MSE_f	Mean observation time, sec
0.5	2.2	3.4	1.4
1.0	2.4	4.0	1.3
2.0	3.6	7.4	1.4

EFFECTS OF TASK DIFFERENCES.—The third experiment was conducted to investigate the effects on two-axis performance of differences between the left- and right-hand tasks. The input bandwidths were either 0.5 rad/sec on both axes, 2.0 rad/sec on both axes, or 0.5 rad/sec on one axis and 2.0 rad/sec on the other. The mean squared inputs (MSI) were either identical on both axes or were adjusted to produce identical one-axis mean-squared error scores (unnormalized) when either input signal was tracked. The controlled-element dynamics were K/s on both axes.

All performance measures were affected by task differences. Table 5 shows that as the difficulty of the task on axis b was increased relative to the difficulty of the task on axis a (in terms of mean-squared error), the subject devoted an increasing share of his foveal attention to axis b and allowed his mean-squared error performance on axis a to deteriorate. The fraction of foveal attention and the fraction of the total two-axis MSE allocated to a given axis were nearly equal on the average. Corresponding decreases in the controller's gain and increases in remnant accompanied increases in mean-squared error. Except for differences in gain, the one-axis foveal and two-axis describing functions were similar. (Peripheral describing functions were not measured in this experiment.)

TABLE 5.—EFFECT OF TASK DIFFERENCES ON
MEAN-SQUARED ERROR AND MONITORING BEHAVIOR
WITH A DISPLAY SEPARATION OF 30°
[Average of three subjects]

Conditions on axis b	MSE_2/MSE_f	Mean observation time, sec	Fractional total MSE	Allocations foveal attention
Axis a=Low-bandwidth axis (BW=0.5 rad/sec) .				
BW=0.5, same MSI	2.2	1.4	0.50	0.50
BW=2, different MSI	2.6	1.3	.45	.43
BW=2, same MSI	8.2	1.2	.39	.37
BW=0.5, same MSI	2.8	2.0	0.61	0.63
BW=0.5, different MSI	3.0	1.7	.55	.57
BW=2, same MSI	3.6	1.4	.50	.50

A SIMPLE MODEL FOR TRACKING WITH SEPARATED DISPLAYS

It appears from our experimental results that multiaxis models for the human controller can be constructed from a simple-combination of single-axis models of the type presented by McRuer et al. (ref. 2). The art of modelling multiaxis control situations will be greatly facilitated if this proves to be true in general. One will then be able to predict system performance from the system parameters by applying a set of straightforward combinatorial rules to the existing human controller models.

A simple switching model of the human controller predicts with reasonable accuracy the effects of visual scanning on system performance. The key assumption of this model is

that the human controller acts as a two-channel processor of information: one channel processes information obtained foveally while the other simultaneously processes information obtained peripherally. There is assumed to be no coupling, or interference, between channels.

These assumptions lead to the model shown in figure 3, in which the human controller's strategy on each axis of a two-axis task can be represented by two dynamic elements whose

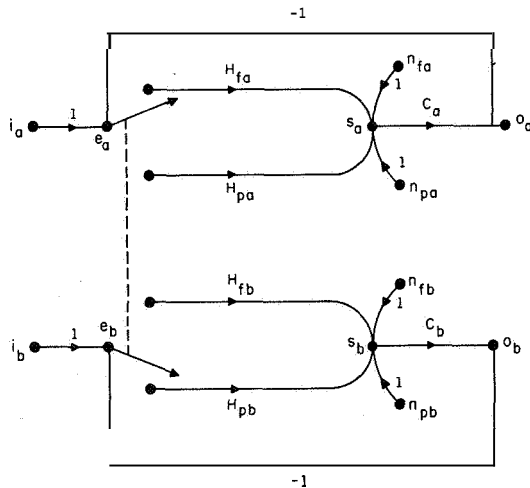


Figure 3.—Model of the two-axis manual control situation with separated displays. Input= i , system error= e , control stick motion= s , system output= o , human controller's remnant= n , human controller's describing function= H , controlled element dynamics= C , and foveal and peripheral strategies are denoted by subscripts f and p .

display separation of 30° . No remnant was simulated other than that due to switching effects.

Simulated one-axis foveal, one-axis peripheral, and two-axis describing functions were obtained from the model. The strategy-selector switch was driven by a periodic waveform unrelated to the forcing function, and the foveal and peripheral dwell times were each 1.25 seconds. The system was driven by a 2 rad/sec waveform of the type used in the manual control experiments.

Figure 4 shows that the three describing functions obtained from this model have identical shapes and differ only in gain. The peripheral gain is 8 dB less than the foveal gain, and the two-axis gain lies halfway between the two. The two-axis model describing function therefore is related to the one-axis describing functions in the same way as the human controller two-axis describing functions are related to the corresponding one-axis foveal and peripheral describing functions (fig. 2). Although there was essentially no remnant associated with the one-axis foveal or one-axis peripheral describing functions obtained from the analog model, the two-axis simulation yielded a fractional "stick" remnant power of 0.21. This is not surprising, since our measurement procedure considered as remnant all signal power not accounted for by a time-invariant linear strategy. We suspect that

outputs are added and whose inputs are switched. One of these elements $H_{fa}(\omega)$ is assumed to be equal to the describing function generated when the subject tracks axis a alone. The other $H_{pa}(\omega)$ is equal to the describing function appropriate to single-axis Peripheral tracking. A second pair of elements describes the controller's strategy on axis b . Remnant terms are associated with all four describing functions. Switching on the two axes is coupled so that $H_{fa}(\omega)$ is applied simultaneously with $H_{pa}(\omega)$. This model does not include sample-and-hold mechanisms. When signal e_a is applied to element H_{fa} , for example, the signal applied to element H_{pa} is assumed to be zero.

A physical representation of the model was constructed on the analog computer so that the model's predictive value could be tested. Experiments were performed to compare the two-axis describing functions and mean-squared error performance of the model with those of the human controllers.

Tracking with K/s dynamics was simulated. The controller's foveal and peripheral describing functions were both simulated by filters of the form $H(s) = K e^{-0.2s}$. The "foveal gain" was adjusted to yield an NMSE as close as possible to that achieved by the human controller when tracking a single axis. The "peripheral gain" was adjusted to yield an NMSE typical of peripheral tracking performance with a

strategy-switching was an important source of the remnant observed in the two-axis manual control experiments.

Another experiment was performed to investigate the relation between the MSE performance of the model and the allocation of "foveal attention." The foveal and peripheral dwell times were varied in such a manner that their sum was 2.5 seconds. The 2 rad/sec waveform was used to drive the system. The following relationship was observed:

$$MSE_{2a} = A \cdot MSE_{fa} + (1-A)MSE_{pa} \quad (2)$$

where A is the fraction of time spent tracking axis a foveally. The applicability of these results to the manual control situation was determined by using the one-axis foveal MSE scores, the one-axis peripheral MSE scores, and the eye-movement data obtained in the manual control experiment to predict the two-axis MSE scores according to equation (2). Table 6 shows that the predicted and measured average two-axis scores generally differed by less than 10 percent. The greatest predictive errors occurred for the control situation in which the input bandwidths were different and the mean squared inputs were the same. The two-axis MSE measured on the low-bandwidth axis was about three times that predicted by equation (2), whereas the MSE observed on the high-bandwidth axis was about 20 percent lower than the predicted score. The total-task MSE, however, was predicted with an error of less than 20 percent.

The model of figure 3 is only a partial model of the multiaxis control situation; it does not provide accurate predictions of scanning behavior. For example, the subjects rarely scanned at an average rate greater than one scanning cycle per 2 seconds, even though the two-axis NMSE score obtained from the analog model decreased continually as the switching rate was increased without limit. Modifications to the model, such as cost weightings for eye movements, are necessary to account for limitations on the human controller's scanning rate.

CONCLUSIONS

We have given a brief review of our studies of two-variable manual control systems. The results obtained suggest that the single-axis describing function models provide a good starting point for predicting the human controller's two-axis performance. These models must be modified to take into account interaction between channels when the dynamics are different in the two axes and the effects of peripheral viewing of the display when the displays are separated. The data we presented for peripheral tracking represent an important addition to the manual control data base and can be used directly to predict two-variable performance.

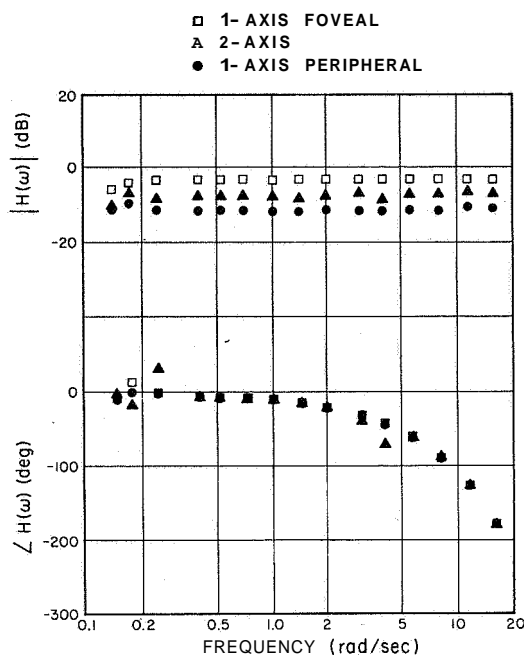


Figure 4.—Model describing functions for one-axis foveal, one-axis peripheral, and two-axis simulated tracking conditions. Controlled element dynamics were K/s on both axes and input bandwidths were 2 rad/sec on both axes. Simulated display separation was 30°, and 0 dB=10⁵ dynes/cm.

TABLE 6.—PREDICTED AND MEASURED TWO-AXIS
MEAN-SQUARED ERRORS SEPARATED DISPLAYS

Experimental variable	Predicted MSE, cm ²	Measured MSE, cm ²
Effects of display separation		
Separation=30°	0.80	0.76
Separation=56°	1.72	1.64
BW=0.5 rad/sec	0.77	0.77
BW=1.0 rad/sec	.77	.72
BW=2.0 rad/sec	.84	.72
Low-bandwidth axis	0.11	0.33
High-bandwidth axis	.64	.52
Total task	.75	.85
Low-bandwidth axis	0.43	0.49
High-bandwidth axis	.72	.61
Total task	1.15	1.10
Dynamics=K	0.52	0.53
Dynamics=K/s	.77	.72
Dynamics=K/s ²	.80	1.00

monitor theory (refs. 3 and 18) must be modified to predict visual sampling behavior in a combined monitoring-control task. The fact that we encountered observation times three or four times larger than those found in pure monitoring tasks suggests that the modifications required may be considerable. The importance of peripheral vision in control tasks also suggests that this factor should be given more consideration than it usually is in most monitoring studies. However, as we go to higher-dimensional tasks, the complexity of the visual field may be so great that the relative contribution of peripheral tracking in any one control loop will be small. If this turns out to be the case, we may be able to simplify the structure of the multi-axis model considerably by considering only foveal monitoring and control behavior.

REFERENCES

1. Elkind, J. I.: Characteristics of Simple Manual Control Systems. TR-11, MIT Lincoln Laboratory, April 6, 1956.
2. McRuer, D. T.; Graham, D.; Krendel, E. S.; and Reissener, W., Jr.: Human Pilot Dynamics in Compensatory Systems—Theory, Models, and Experiments With Controlled Element and Forcing Function Variations. AFFDL-TR-65-15, July 1965.
3. Senders, J. W.: The Human Operator as a Monitor and Controller of Multidegree of Freedom S-Systems. IEEE Transactions on Human Factors in Electronics, vol. HFE-5, no. 1, Sept. 1964, pp. 2-5.
4. Fitts, P. M.; and Simon, C. W.: Effect of Horizontal Vs Vertical Stimulus Separation on Performance in a Dual Pursuit Task. Amer. Psychologist, vol. 4, 1949, pp. 304-305.
5. Chernikoff, R.; et al.: Two Dimensional Tracking With Identical and Different Control Dynamics in Each Coordinate. NRL Rep. 5424, Nov. 27, 1959.
6. Duey, J. W.; and Chernikoff, R.: The Use of Quickening in One Coordinate of a Two-Dimensional Tracking System. IEEE Transactions on Human Factors in Electronics, HFE-1, Mar. 1960, pp. 21-24.
7. Chernikoff, R.; and LeMay, M.: Effect of Various Display-Control Configurations on Tracking With Identical and Different Coordinate Dynamics. J. Exper. Psych., vol. 66, no. 1, 1963, pp. 95-99.
8. Verdi, A. P.; et al.: Effects of Display Quickening on Human Transfer Functions During a Dual-Axis Compensatory Tracking Task. ASD-TDR-65, Wright-Patterson Air Force Base, Mar. 1965.
9. Todosiev, E. P.; Rose, R. E.; Bekey, G. A.; and Williams, H. L.: Human Tracking Performance in Uncoupled and Coupled Two-Axis Systems. TRW Systems, Rep. 4380-6003-R0000, Dec. 8, 1965.
10. Levison, W. H.; and Elkind, J. I.: Studies of Multi-Variable Manual Control Systems: Two-Axis Compensatory Systems With Compatible Integrated Display and Control, Rep. no. 1339, Bolt Beranek and Newman, Inc., 1966.
11. Wierwille, W. W.; and Gagne, G. A.: Nonlinear and Time-Varying Dynamical Models of Human Operators in Manual Control Systems. Human Factors, vol. 8, Apr. 1966, pp. 97-120.
12. Sanders, A. F.: The Selective Process in the Functional Visual Field. Van Nostrand Reinhold, N. Y., Amsterdam, Netherlands, 1963.
13. Senders, J. W.; Webb, E. B.; and Baker, C. A.: The Peripheral Viewing of Dials, Appl. Psychol., vol. 39, no. 6, 1955.
14. McCollgin, F. H.: Movement Thresholds In Peripheral Vision. J. Optical Soc. Am., vol. 50, Aug. 1960, pp. 774-779.
15. Graham, C. H. (ed.): Vision and Visual Perception. John Wiley & Sons, Inc., 1965.

16. Grignetti, M. C.: Design of a Digital Computer Facility for Flight Simulation Experiments, Third Quarterly Progress Report on Job. No. 11206, Bolt Beranek and Newman, Inc., Apr. 1966.
17. Elkind, J. I.; Starr, E. A.; Green, D. M.; and Darley, L. D.: Evaluation of a Technique for Determining Time-Invariant Dynamic Characteristics of Human Pilots. NASA TN D-1897, 1963.
18. Senders, J. W.; Elkind, J. I.; Grignetti, M. C.; and Smallwood, R. D.: An Investigation of the Visual Sampling Behavior of Human Observers. NASA CR-434, Apr. 1966, pp. 27-33.

3. Piloted Simulator Display System Evaluation — Effective Resolution and Pilot Performance in the Landing Approach

*Wendell D. Chase
Ames Research Center, NASA*

A study was conducted in two parts to investigate the quality of a visual display in a fixed cockpit piloted simulator and ways of measuring pilot-vehicle performance. Part I concerned the effective resolution of a typical simulator display relative to that for the real world; part II concerned pilots' estimates of range and altitude from the runway threshold and measures of his ability to control the vehicle in the approach and landing. A correlation analysis of the information from part II was used to indicate the degree of association between those performance measures.

The static display characteristics, as measured by the resolution of landolt C-rings, were found to be degraded by as much as a factor of 12 when compared with the real world. A further loss of resolution by approximately one-third of the static resolution occurred with the moving display and was influenced by the apparent motion of the airplane.

Range estimates to the runway threshold were in error by about 10 percent; altitude estimates above the runway threshold were in error by about 20 percent. Error in range estimates decreased with experience while altitude estimates remained relatively constant.

Performance in the landing approach was very similar to that in actual flight and even included a "duck under" maneuver by each pilot. The termination of the landing approach was at higher rates of descent, but touchdown distance from the runway threshold was about the same as that in actual flight.

A correlation analysis between the various measures of altitude-range estimates, and pilot-vehicle landing performance showed the following: (1) that the touchdown error depends on the pilot's ability to judge altitude in the landing approach, and (2) the touchdown error is highly correlated with the integrated altitude error, and the correlation indicates difficulty in estimating the correct altitude to decrease the rate of descent and to initiate the flare. However, the absence of motion feedback, ground effect dynamic forces, and vestibular and kinesthetic cues may be partially responsible for these errors.

Piloted simulators are being used extensively for research related to advanced aircraft and spacecraft. An important component of some of these simulators is a display of the outside scene for the crew. Because of limitations in state-of-the-art electronics and optics, these visual display systems do not provide a true picture of the real world. Although considerable progress has been made in recent years in developing display systems for research and training (refs. 1 and 2), quantitative studies of the visual quality of these systems are needed to relate the display characteristics with man-system performance. This information could provide a more rational basis for defining requirements and specifications for

piloted-simulator display systems. Although it was recognized that correlates between basic characteristics of visual display systems (e.g., resolution and pilot-vehicle performance measures) would be difficult to establish because of pilots' adaptive capabilities, it was considered desirable to study an available display system to provide some baseline information in this area.

Accordingly, the present studies had the following primary objectives:

(1) To define experimentally the effective static resolution of a television display relative to that for the real world, and to measure the loss of resolution that results from motion of the aircraft at landing-approach speeds.

(2) To determine pilots' ability to estimate range and altitude with a simulator display system and to land a representative commercial jet transport.

In the first part of this paper, the effective resolution characteristics of the simulator display system are presented for several pilots and compared with those for the real world. In the second part of the report, variations in performance of several pilots for several performance measures, obtained in landing approaches, are provided, and some tentative correlations among these pilot performance measures are briefly noted.

SYMBOLS

f	number of favorable ways
\bar{h}	mean altitude, ft
$\dot{\bar{h}}$	mean rate of descent, ft/sec
N	event of different ways
P	probability
r	correlation, dimensionless
\bar{S}	mean touchdown distance from runway threshold, ft
$ \bar{S}_e $	mean absolute touchdown error from glide-slope runway intersection, ft
\bar{S}_R	mean resolvable distance of C-ring, ft
\bar{V}	mean touchdown airspeed, knots
β	mean resolvable angle of C-ring, min
$\bar{\sigma}_e$	mean standard deviation error at discrete altitudes, ft
σ_{eh}	standard deviation of altitude estimates, dimensionless
σ_{er}	standard deviation of range estimates, dimensionless
σ_h	standard deviation of \bar{h} , ft
σ_S	standard deviation of \bar{S}_R , ft

σ_{sd} standard deviation of \bar{S} , ft

σ_v standard deviation of \bar{V} , knots

$\bar{\int} h_e$ mean integrated altitude error between command and actual flight path, ft

EQUIPMENT AND METHOD

DESCRIPTION OF APPARATUS.—The components of the visual simulator are a television camera, runway model (scaled at 300 ft=1 ft), projection system, and cab. The television camera mounted on a five-degrees-of-freedom carriage assembly is shown in figure 1. The belt (runway) transporter with one degree of freedom is also shown. The television camera is a General Electric 525 scan line, 30 frames/sec, 2:1 interlace, 4:3 aspect-ratio system. The front projection system is a Schmidt projector, with correction plates and a retro-reflective screen with a gain of 2.5. The field of view afforded the pilot, located 10 feet from the screen, was 50° horizontal and 37.5° vertical. An Elgeet 13-mm, f/2.5 wide-angle lens was used to produce a unity magnification ratio.

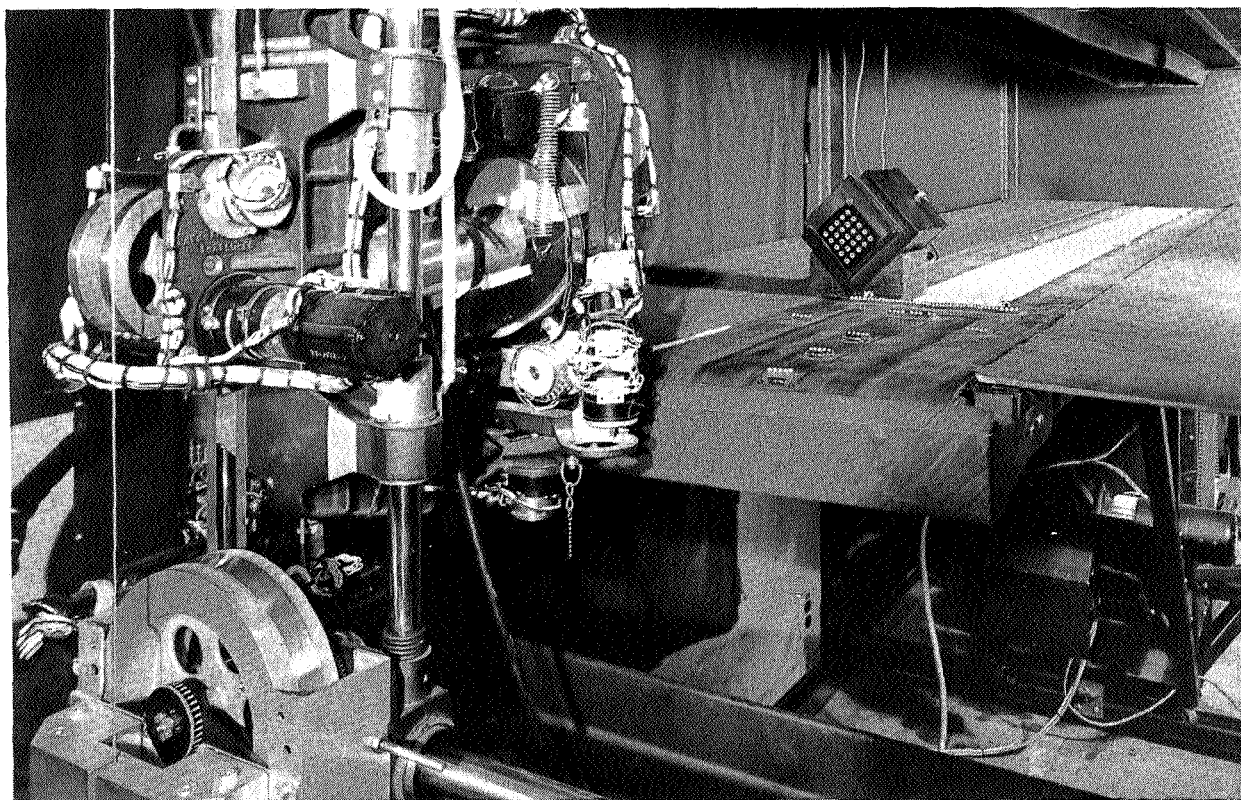


Figure 1.—Landing-approach model assembly.

EXPERIMENTAL PROCEDURES.—Three pilots participated: one Ames test pilot and two engineers with military flight experience. Simulator resolution was investigated by measuring the pilots' performance as they viewed stationary or moving objects. The

performance index used was essentially the pilot's visual acuity for a test object observed under a static or dynamic condition.

Resolution Characteristics (static).—A special light box and matrix plate (fig. 2) was constructed for the pilots to observe the orientation of landolt C-rings, that is, left, down, right, or up. This method, which has no form discrimination, was designed to determine the relative acuity of the observers as a function of the television resolution. The average light reading of the matrix box that minimized blooming at the projector was determined to be 66.4 foot-lamberts with a contrast ratio of 99.2 percent. The complete matrix box was rotated 45° (figs. 1 and 2) because of an unequal horizontal and vertical scan of the television camera. All C-rings within the matrix thus have the same number of picture elements transmitted by the vidicon. Each individual C-ring was scaled at 6 feet in diameter from which the mean resolvable angle can be determined.

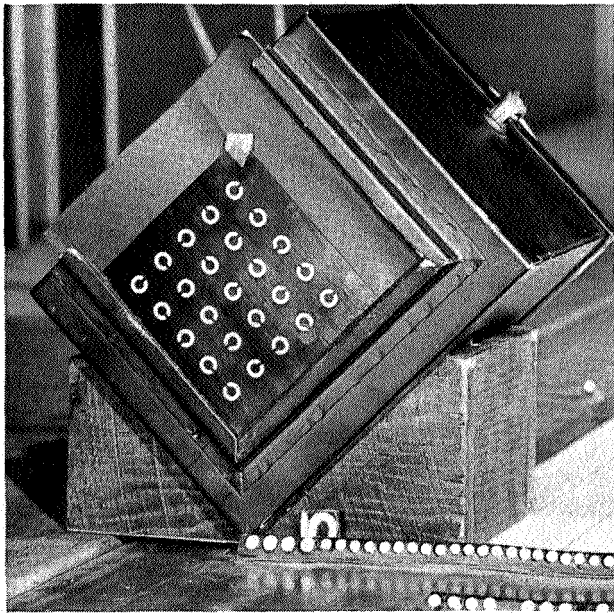


Figure 2.—Light source and landolt C-ring matrix.

Resolution Characteristics (dynamic).—In order to assess pilot performance, on the basis of information used for visual contact with the runway threshold during aircraft landings, it was necessary to produce relative measurements comparable to observations by human subjects with normal viewing conditions. This was accomplished by using a single, landolt C-ring rotated from run to run, and located at a comparable real-world wheel height above the runway threshold. The pilot's task was to acquire the runway threshold visually during a normal landing approach and, when the orientation of the ring became discernible, to activate a switch that would record altitude, range to threshold, speed, and other pertinent variables.

Range and Altitude Estimates.—The pilot's ability to estimate altitude and range, at discrete points along a normal 3° glide slope to the runway threshold, was measured by normalized error responses. Altitude varied from a minimum of 30 feet to a maximum of 270 feet, while the range varied from a minimum of 500 feet to a maximum of 2500 feet.

The experimental design required a pilot-run matrix of 25 altitude-range combinations that were displayed with equal probabilities in seven variations according to a Latin square experimental design.

Landing-Approach Performance.—Pilot performance was measured in actual landing approaches and in a fixed cockpit landing-approach simulator (fig. 3) with the Boeing 707 dynamics. The principal dynamics used were those for the aircraft longitudinal response including both the phugoid and short period. The pilot's approach task was to establish a stable, well-controlled rate of descent to the runway threshold with visual references and to



Figure 3.—Pilot station and projected runway image.

make a termination maneuver including a successful flare and touchdown. Initially, the altitude was 300 feet, the flight path was 3° , the ground distance to runway threshold was 5731 feet, and the approach airspeed was 135 knots. Each pilot made a total of 45 approaches and landings.

RESULTS AND DISCUSSION

DISPLAY RESOLUTION.

Static: 100-Percent Correct Detection of C-Ring Opening.—The results of 25 observations per pilot of the static landolt C-ring matrix through the television system are summarized in table 1. Of primary interest are the mean resolvable distance s , standard deviation σ_s , mean resolvable angle θ , and Snellen acuity for each pilot.

The mean resolvable angle θ , subtended at the eye by the C-ring opening, is a measure of the minimum separable acuity which, in this case, is 4.79 minutes of arc or 287.4 seconds of arc. The nominal value resolved by human subjects under normal viewing conditions is 24 seconds of arc (ref. 3). Thus, this is a reduction in resolution equivalent to a reduction

Parameter	Pilot A	Pilot B	Pilot C
\bar{S}_R , ft -----	846.48	869.86	862.82
σ_S , ft -----	13.67	3.05	5.29
β , min -----	4.87	4.74	4.78
Snellen acuity ----	20/13	20/13	20/20

$$P=f/N$$

where :

f number of correct ways
N number of different ways

Let

A a correct answer
 B subject detecting ring position
 C correct answer by guessing only

then

$P(C)$ 1/4 (since there are four possible positions of the C-ring)
 $P(\bar{C})$ 3/4
 $P(B)$ 1/2 at threshold (50 percent detected correctly)
 $P(\bar{B})$ 1/2 at threshold (50 percent not detected correctly)

Therefore,

$$P(A)=P(B)+P(\bar{B})P(C)=1/2+1/2 \cdot 1/4=5/8$$

The probability of a wrong answer is

$$P(\bar{A})=P(\bar{B})P(\bar{C})=1/2 \cdot 3/4=3/8$$

Therefore, in 25 tries the expected number of correct responses at the subject's sensitivity threshold is

$$(25)P(A)=25(5/8)=15 \frac{5}{8}$$

Thus, the just resolvable threshold distance can be determined if each pilot can correctly identify 16 landolt C-ring positions consistently out of the 25. The results show that approximately 100 feet can be added to each pilot's mean resolvable distance \bar{S}_R for each respective threshold distance. Consequently the televisual resolution ratio indicates a reduction in acuity of about 11:1.

Dynamic.—The dynamic flight landing approach to the runway threshold required the pilots to observe the orientation of the landolt C-ring as it was rotated randomly, but with the same order for each pilot, from landing approach to landing approach. Table 2 summarizes the parameters of mean range \bar{S}_R , mean altitude \bar{h} , their respective standard deviations σ_S and σ_h , and mean resolvable angle β for which acuity is maximum.

TABLE 2.—LANDING-APPROACH C RING SUMMARIES

Parameter	Pilot A	Pilot B	Pilot C
\bar{S}_R , ft -----	601.47	565.04	575.86
σ_S , ft -----	47.68	54.22	65.95
\bar{h} , ft -----	93.08	76.70	89.08
σ_h , ft -----	6.02	6.40	6.28
β , min -----	6.86	7.30	7.16

The dynamic mean resolvable angle β for all pilots if 7.10 minutes of arc are converted to minimum separable acuity is 426 seconds of arc. The ratio between static and dynamic minimum separable acuity is

$$\frac{\text{'static}}{\text{'dynamic}} = \frac{287.4}{426} = 0.6746$$

This figure is analogous to the pilot's ability to discriminate visual detail of a moving object and is sometimes called dynamic visual acuity (refs. 4 and 5). The change in acuity of approximately one-third under simulated pilot approach dynamic conditions, which could be expected in the real-world comparison, shows that the loss in resolving power is strongly influenced by the apparent motion of the airplane. It is quite possible that this ratio could be further impaired for aircraft with higher approach speeds.

It is interesting to note from figure 4 that pilot A, whose static acuity was the worst, had a better dynamic acuity than pilots B and C. Likewise, pilot B, whose static acuity was the best among the pilots, had the worst dynamic acuity. Pilot C falls between pilots A and B for both static and dynamic acuity. The significance is that the pilot's dynamic and static acuity are not the same. This difference may be influenced by each pilot's training and his particular method of controlling the aircraft in the landing approach.

PILOT PERFORMANCE MEASURES

Range and Altitude Estimation.—The pilot performance in estimating range and altitude are represented in figure 5. The standard deviations σ_{eh} and σ_{er} of errors are shown as a relative error since the pilot's response to actual altitude and range has been normalized with respect to the actual altitude. Each pilot shows an asymptotic level in estimating range after at least five sessions. This indicates that some learning has taken place. Little learning is indicated for altitude estimates except for pilot C. Generally, all pilots estimated range twice as well as they estimated altitude, which may indicate more horizontal visual cues than vertical cues. Although the pilots attempted to judge altitude by the intersection of the runway with the horizon, their relative altitude estimate errors still exceeded their range estimate errors.

The mean standard deviation $\bar{\sigma}_e$ in feet at discrete altitudes is shown in figure 6 for all pilots. The figure indicates that errors in altitude estimates are around 15 percent and not 20 percent as indicated from the pilot performance curves of figure 5, which were the overall responses.

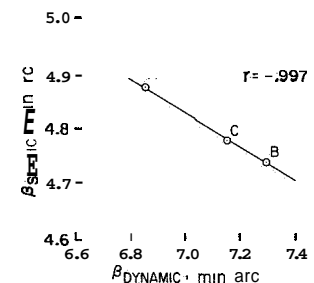


Figure 4.—Pilots' static and dynamic resolving angle.

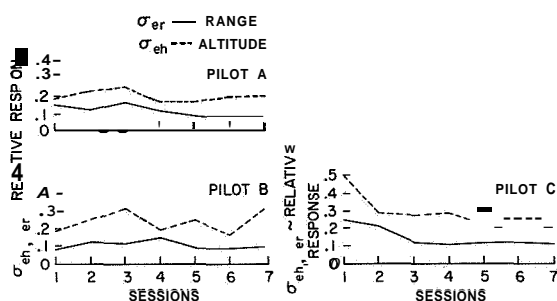


Figure 5.—Pilot performance curves for estimates of range and altitude.

tance errors S_e over the span of the trials for pilots A and C have a mean error less than 400 feet, which compares favorably with real flight landings (ref. 6); however, pilot B has a greater mean error of about 700 feet.

The lines of regression for figure 7 show trends over wide variations. Appreciable negative or positive regression line slopes (\pm correlation) indicate either an increment or decrement in learning. However, since the slopes for all three pilots are small, learning does not appear to be a factor in this particular performance measure.

The integral absolute altitude error is a measure of how well the pilot adheres to the command flight path until touchdown. Figure 8 indicates a mean integrated altitude error of about 200 feet for pilot A, about 300 feet for pilot B, and only about 160 feet for pilot C. The larger errors of pilots A and B are related to a "duck-under" maneuver (ref. 6) executed upon visually acquiring the runway threshold.

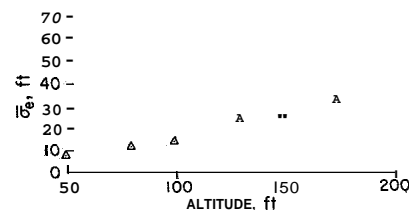


Figure 6.—Pilots' average altitude error at discrete altitudes.

Parameter	Pilot A	Pilot B	Pilot C
\bar{b} , ft/sec ---	4.51	3.72	2.77
$\sigma_{\bar{b}}$, ft/sec --	1.75	1.50	1.56
\bar{S} , ft -----	1703.78	1368.49	1841.53
$\sigma_{\bar{S}}$, ft ----	236.66	451.30	361.51
\bar{V} , knots ---	109.65	109.32	124.03
$\sigma_{\bar{V}}$, knots---	4.32	5.41	3.17

Table 3 shows that the simulator mean rate of descent is higher (2.77 ft/sec to 4.51 ft/sec) than that recorded in real-flight visual flight rules (VFR) conditions of about 2 ft/sec (ref. 6). The problem associated with higher rates of descent for the simulated landings indicates that the pilots had some difficulty in estimating altitude prior to starting the flare. Although the "ground effect" equations were not included in the simulation, there may be

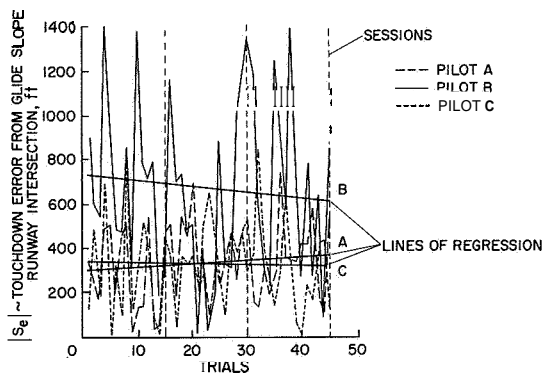


Figure 7.—Pilots' touchdown error versus landing trials.

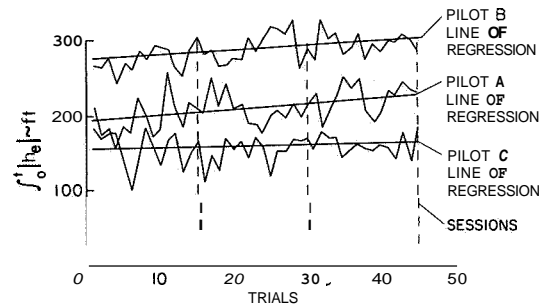


Figure 8.—Pilots' integrated altitude error versus landing trials.

other effects which degrade performance such as missing motion feedback, vestibular and kinesthetic cues, or limited image resolution (ref. 7).

Mean distance traversed from the runway threshold during VFR conditions is about 1500 feet for real flight (ref. 6), which is comparable with the piloted simulator range (subject means) of 1370 to 1840 feet. Wheel height above threshold recommended by the International Civil Aviation Organization (ICAO) standard is 40 feet; flight data from 108 landings at Kennedy International show a predominate wheel height of 20 feet (ref. 6); the simulator data show a wheel height of about 35 feet.

Performance Correlates.—Correlation is simply the similarity, in direction and degree, of variations in corresponding pairs of observations of two variables. The principal problem of simple correlation is that of determining the degree of association between these pairs of observations. The aggregate of plotted points was consolidated into averages (means) and standard deviations in order to investigate more easily the pilots' scatter diagram that might show a trend. The pilot scatter diagram can be expressed by an equation of the trend called "the line of regression," which is a minimum-squared-error linear curve fitted to the scatter diagram. The slope of this line depends upon the coefficient of correlation r , $-1 \leq r \leq 1$. It appears that the closer the points lie to a line of regression, the more nearly a simple linear equation expresses the association between the variables. Thus, a few measured parameters were thought to contribute to problems associated with the landing approach and were tested for a correlation coefficient close to ± 1 (indicating a definite linear relationship between the variables). Those with a correlation coefficient close to zero (indicating practically no linear relationship) were not included in this report.

Figure 9 shows the relationship of the standard deviation of velocity σ_v versus the mean integral altitude error $\bar{\int} h_e$ to be highly correlated ($r=0.983$).

Figure 10 shows the mean absolute touchdown error $|\bar{s}_e|$ versus the mean integral altitude error with a high correlation ($r=0.94$).

Figure 11 shows the mean integral altitude error $\bar{\int} h_e$ versus the mean touchdown distance from the runway threshold \bar{S} to be highly correlated ($r=0.995$). This high correlation indicates difficulty in estimating the correct altitude to level off the rate of descent prior to and at the start of the flare.

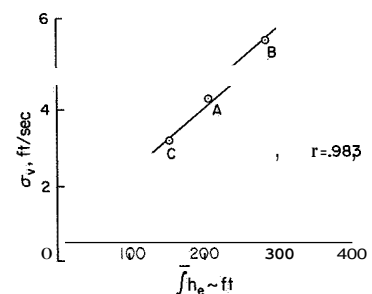


Figure 9.—Pilot correlation between velocity and integrated altitude error.

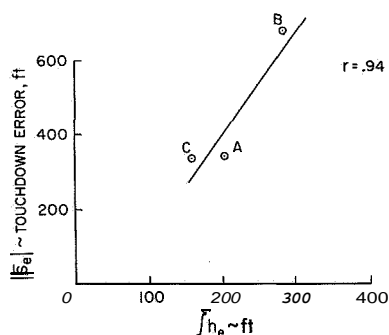


Figure 10.—Pilot correlation between integrated altitude error and touchdown error.

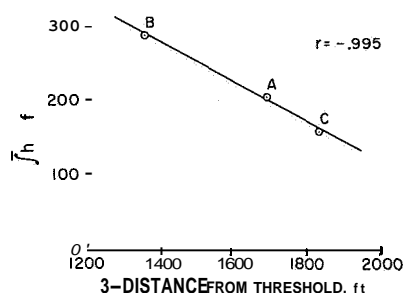


Figure 11.—Pilot correlation between touchdown distance and integrated altitude error.

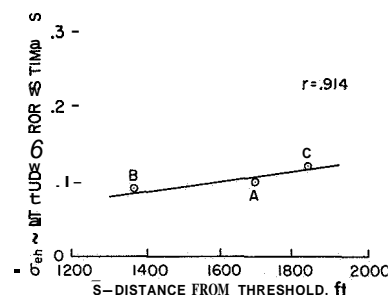


Figure 12.—Pilot correlation between touchdown distance and static altitude error estimates.

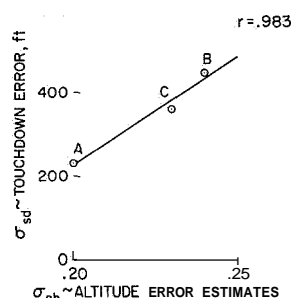


Figure 13.—Pilot correlation between standard deviations of static altitude error estimates and touchdown error.

The relationship between the pilots' static standard deviation of altitude estimates σ_{eh} and the simulated flight mean touchdown errors \bar{S} is shown in figure 12 to be highly correlated ($r=0.914$). This may be significant in that the mean distance from runway threshold may be due chiefly to the pilots' inability to judge altitude correctly in the landing approach.

Similarly, figure 13 shows the standard deviation of touchdown error σ_{sd} versus the standard deviation of static altitude estimates σ_{eh} to be highly correlated ($r=0.983$). This shows that the touchdown error is closely related to the pilots' inability to judge altitude in the landing approach.

CONCLUSIONS

The resolution characteristics of the display system, determined from measurements of static acuity, show a considerably degraded resolution equivalent to a reduction in acuity of 12 to 1. An additional loss of resolution (dynamic visual acuity) by approximately one-third occurred in the landing approach.

Static and dynamic visual acuity apparently have a negative correlation. Pilots have different static acuity but dynamic acuity may be further influenced by pilot training and methods of controlling the aircraft in the landing approach. This may account for a negative correlation; furthermore, it shows that the relative motion of the airplane can cause a loss of resolution during a landing approach.

A television-projected simulator display appears to lack the sharpness and clarity of actual flight conditions for the landing approach visual flight rules (VFR) maneuver.

The pilot uses visual cues in both the horizontal and vertical planes to estimate range and altitude. The pilots of the simulator visually estimated altitude with an error of about 20 percent. There was a correlation between altitude estimates and touchdown errors. Range estimate errors were about 10 percent, but there was little correlation between range estimates and touchdown errors.

Real flight touchdown distance (from the threshold) and rate of descent during VFR conditions are somewhat less than those recorded from the piloted simulator. However, these

errors, which degrade performance in the simulator, may be due to the absence of kinematic feedback, vestibular cues, kinesthetic cues, and image degradation.

The correlation analysis of the landing-approach performance has shown the following: (1) the touchdown error is closely associated with the pilots' inability to judge altitude in the landing approach, and (2) the touchdown error is also closely related to the integrated altitude error which indicates the pilot's difficulty in estimating the correct altitude for leveling off the rate of descent prior to and while initiating the flare.

REFERENCES

1. Buddenhagen, T. F.; Johnson, A. B.; Stephan, S. C.; and Wolpin, M. P.: Development of Visual Simulation Techniques for Astronautical Flight Training. Tech. Doc. Rep. AMRL-TDR-63-54, vols. I and II, June and Nov. 1963.
2. Harshbarger, John H.; and Gill, Arthur T.: Development of Techniques for Evaluation of Visual Simulation Equipment. Aerospace Med. Lab., Aug. 1964.
3. Bartley, Samuel Howard: Vision - A Study of Its Basis. Washington Univ. School of Medicine, 1963.
4. Bhatia, Balraj; and Verghese, C. A.: Threshold Size of a Moving Object as a Function of Its Speed. J. Optical Soc. Am., vol. 54, no. 7, July 1964, pp. 948-950.
5. Snyder, H. L.; and Greening, C. P.: The Effect of Direction and Velocity of Relative Motion Upon Dynamic Visual Acuity. Autonetics Human Factors Department C5.447/3111, Jan. 1965.
6. Anon.: The 100-Foot Barrier, Part II. Proposal J-5740A, Cutler-Hammer Airborne Instruments Laboratory, Mar. 1965.
7. Watson, D.: The Calculation of Televisual Detection Range. Tech. Note WE 21, Royal Aircraft Establishment (Farnborough), Apr. 1963.

BIBLIOGRAPHY

- Anon.: Image Evaluation Techniques. Inst. Optics, College of Engineering and Applied Sciences, Univ. Rochester, 1963.
- Baldwin, M. W.: The Subjective Sharpness of Simulated Television Images. Proc. IRE, vol. 28, 1940, pp. 458-468.
- Brock, G. C.; Myskowski, E. P.; and Attaya, W. L.: Study of Image-Evaluation Techniques. Interim Eng. Rep. 5, Itek Corp., Aug.-Dec. 1963.
- Elias, Merrill F.: Speed of Identification of Televised Symbols as a Function of Vertical Resolution. Final Rep. RADC-TR-65-239, July 1965.
- Elias, Merrill F.; Snadowsky, Alvin M.; and Rizey, Edward F.: The Relation of Number of Scan Lines Per Symbol Height to Recognition of Televised Alphanumerics. RADC-TDR-64-433, Oct. 1964.
- Fink, Donald G.: Television Engineering. Second ed., McGraw-Hill Book Co., Inc., 1952.
- Pinsker, W. J. G.: Features of Large Transport Aircraft Affecting Control During Approach and Landing. Presented to Flight Mechanics Panel Take Off and Landing Meeting of AGARD (Paris), Jan. 14-18, 1963.
- Wulfeck, Joseph W.; Weisz, Alexander; and Raben, Margaret W.: Vision in Military Aviation. WADC Tech. Rep. 58-399, Nov. 1958.

4. Holographic Display Systems

A. J. Devaney, C. Grauling, and S. Baron
Electronics Research Center, NASA

Recent advances in the field of holography, particularly in the development of techniques for making high-quality holograms, have raised the possibility of using holograms for flight displays or flight simulators. Holograms possess many characteristics, such as the three-dimensional image formation property and variable perspective, which make them particularly attractive for display or simulation purposes. This is especially true when one wishes to display real scene information since the reconstructed hologram contains almost all the information in the original scene.

A hologram, however, stores a static scene. Thus, if a dynamic display is required, as would be the case if the observer is moving relative to the scene, one is confronted with the possibility that a great many holograms, corresponding to various viewing positions, may have to be stored. An alternative approach would be to manipulate the virtual image resulting from a single hologram in a manner related to changes in viewing aspect and position of the observer. This latter approach for obtaining a dynamic holographic display is the major topic of this paper. In particular, techniques for obtaining six degrees of viewing freedom are developed, using fundamental concepts of optics and holography. These techniques make it possible to display the visual effect of arbitrary motion of the observer relative to the scene using one, or possibly a few, stored holograms.

Some of the current state-of-the-art difficulties associated with holographic display systems are also discussed, and areas for future development are indicated.

FUNDAMENTAL CONCEPTS OF HOLOGRAPHY AND OPTICS

In order to understand properly the approach we have taken to obtain a dynamic holographic display and the advantages and limitations of such a display, it is useful to review briefly some of the fundamental aspects of holography and optics. These ideas can be developed either from a physical or geometric optics point of view (e.g., for a geometric interpretation of holography see ref. 1). We have taken, primarily, the physical optics approach since this approach is readily understood by those with an electronic engineering or related background. However, we shall resort to geometric optics arguments when they serve to clarify the material.

HOLOGRAPHY.—Reference 1 states, "Holography is the art of freezing a light wave into a photographic emulsion so that it can be revived by another light wave." Although the basic principles of holography were well known for some time (ref. 2), real progress awaited the development of a suitable monochromatic, coherent light source, the laser. With the appearance of the laser it became possible to store, on a two-dimensional photographic plate, almost all the visual properties of an object illuminated by the laser. This photographic plate is prepared in a special, but relatively simple, manner and is called a hologram. (The process is called hologram construction.) When the hologram is illuminated by a laser (reconstruction), the observer sees an image of the original scene as viewed through a "window" the size

of the hologram.¹ This image is three dimensional, changes perspective with a change in viewing position and exhibits parallax between near and far objects in the image. To an observer the image would be virtually indistinguishable from the original scene.

The mathematics of hologram construction and reconstruction is not particularly difficult. Consider the procedure for constructing a hologram illustrated in figure 1. The field $E_H(x, y)$ which exists over the aperture of the hologram, is simply the sum of the field $E(x, y)$ produced by the light scattered from the object and the field $E_C(x, y)$ produced by the constructing reference light source. The photographic plate, which when exposed in this fashion is called the hologram, records the intensity² of the field E_H . Thus, when developed, the plate has an optical transmittance given by

$$\begin{aligned} T &= |E_H|^2 = |E + E_C|^2 = (E + E_C)(E + E_C)^* \\ &= |E|^2 + |E_C|^2 + E^* E + E E_C^* \end{aligned} \quad (1)$$

where $|E|$ denotes the magnitude of E and asterisks denote complex conjugates. We have omitted explicit notation of the arguments of functions.

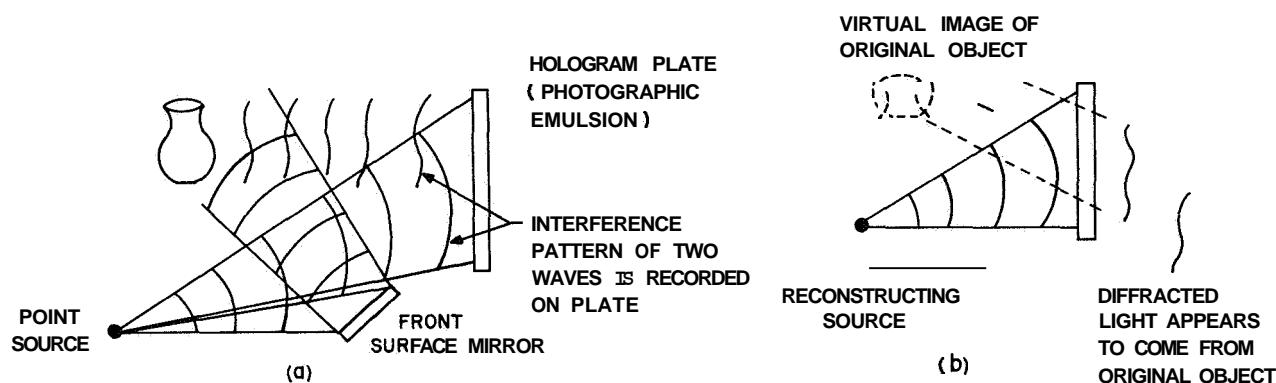


Figure 1.—The hologram. (a) Construction, (b) Reconstruction.

From equation (1) we see that, by the simple expedient of interfering light scattered from the object with a reference beam, we are able to record on a photographic plate a term $E_C^* E$ which is directly proportional to the field produced by the original scene. Herein lies the "secret" of holography for it is just this capability which allows us to reconstruct an image which is a true replica of the original scene. To see this most simply, let us assume that the hologram is reconstructed with a wave E_R which is identical to the original constructing reference wave and, further, that this wave has unit magnitude. Then the field, E_H , produced over the aperture of the hologram is (ref. 6)

¹It is now possible to reconstruct holograms using incoherent light (refs. 3, 4, and 5).

²Actually, some power of the intensity is recorded. This power depends on the characteristics of the photographic emulsion but it can be assumed to be unity for most cases.

³Another way of looking at this procedure is to note that by this procedure we record both amplitude and phase information concerning the original scene. In ordinary photography only amplitude information is recorded.

$$E_H = TE_R = TE_c = \left(|E|^2 + |E_c|^2 \right) E_c + E + E_c^2 E^* \quad (2)$$

Thus, we see that the field over hologram contains the desired original field. There are, however, in equation (2) additional and, in some sense, unwanted terms. We shall discuss these terms momentarily. For now, let us assume that the "unwanted" terms can be suitably attenuated. Then, the observer when looking at the hologram sees a reconstruction of the original field produced by the object; that is, he sees a virtual image which is identical to the original scene. (This image is called the first-order virtual image.) Of course, since the hologram is of finite dimensions, he sees this image through a "window" the size of the hologram.

Now let us briefly examine the remaining terms in equation (2). The first term, called, the zero-order image, is simply the field corresponding to the illuminating wave E_c multiplied by a constant and, as such, contains no useful information. The last term is called the first-order real image or the "twin wave." The "twin" wave is of no immediate interest to us except that we must, effectively, remove it. Now, it can be shown (ref. 7) that the three terms in equation (2), corresponding to three different waves, are propagated in different directions (fig. 2). The "trick" then in making high quality holograms is to separate sufficiently the direction of propagation of the desired reconstructed wave from that of the "twin" wave. This can be accomplished by a suitable skewing of the constructing reference beam (ref. 8). Such skewing of the reference beam was not feasible until the advent of the laser.

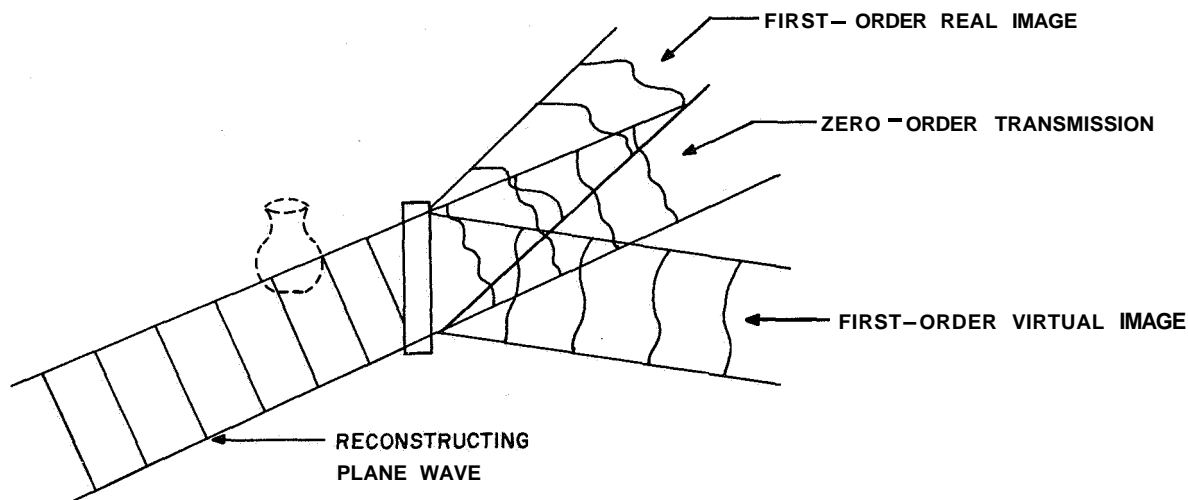


Figure 2.—Direction of propagation of the three waves resulting from a reconstructed hologram.

OPTICS.—The hologram image manipulation techniques which will be discussed can be treated as if one were viewing the scene through a system containing simple optical elements? The elements which we will consider are the thin prism and the negative, or concave, thin

⁴The image quality of holograms has been studied using the analogies with simple optical elements. For example, aberrations and imaging properties of holograms have been compared to those of lens systems (ref. 9). However, these techniques have not been directly applied to image manipulation for dynamic display.

lens. As the name implies, these elements are such that their thickness may be neglected. The operation of such elements may be characterized by a transmission function (ref. 6).

This function is simply the ratio of the output field to the incident field for the given element. The transmission functions for the elements of interest are

$$\text{Thin prism: } E_o/E_i = \exp j \frac{2\pi}{\lambda} \alpha x \quad (3)$$

$$\text{Thin negative lens: } E_o/E_i = \exp -j \frac{\pi}{\lambda f} (x^2 + y^2) \quad (4)$$

where λ is the wavelength of the incident light, α is the sine of the prism angle, f is the focal length of the lens, and the z -axis is chosen coincident with the optical axis of the element.

For reasons which will soon become apparent we will now consider the image-forming properties of the simple lens-prism system of figure 3. The transmission function for this system is

$$E_o/E_i = \exp j \frac{2\pi}{\lambda} \left[\alpha_1 x + \alpha_2 y - \frac{(x^2 + y^2)}{2f} \right]$$

Now, assume that E_i is the field resulting from a point source located at $(\hat{x}, \hat{y}, \hat{z})$, that is,

$$E_i = \exp -j \frac{2\pi}{\lambda} \left[\frac{(x - \hat{x})^2 + (y - \hat{y})^2}{2\hat{z}} \right]$$

After considerable manipulation, one obtains for the output (image) field:

$$E_o = \exp -j \frac{2\pi}{\lambda} \left[\frac{(x - \xi)^2 + (y - \eta)^2}{2\rho} \right] \exp j\phi \quad (5)$$

where ϕ is a constant phase factor and

$$\xi = \hat{f} (\hat{x} + \alpha_1 \hat{z})$$

$$\eta = \hat{f} (\hat{y} + \alpha_2 \hat{z})$$

$$\rho = \hat{f} \hat{z}$$

$$\hat{f} = \frac{f}{f + \hat{z}}$$

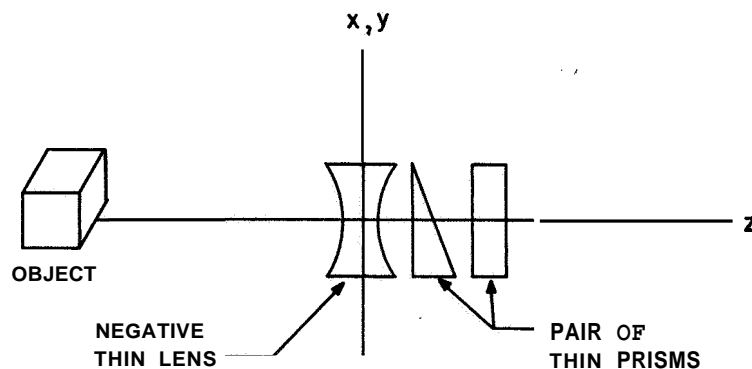


Figure 3.—Simple lens prism system.

Thus, E_0 is simply the field due to a point source at (ξ, η, ρ) modified by a constant phase factor which does not affect the image quality. Hence, each point of an object, when viewed through the lens-prism system, is imaged at a point given by equation (5). To put it another way, the lens-prism system maps object points into image points with the mapping defined by equation (5).

If no prism is present in the optical system, we see from equation (5) that the image will appear smaller than the object and closer to the observer (since $f \leq 1$). (This, of course, is the well known imaging property of a concave lens.) The result in introducing the prisms is an apparent rotation of the object through angles $\alpha_1 \hat{f}$ and $\alpha_2 \hat{f}$ about the y- and x-axes, respectively. The final image produced by the lens prism system is a combination of the separate effects of each element. Thus, to the observer looking through the system, the object appears smaller, closer, and rotated.

The image resulting from such a system is not free from distortion. In particular, \hat{f} depends on the z coordinate of the object point and, consequently, the demagnification and range change will vary for object points with different z coordinates. Also, the prisms actually introduce a "shear" effect rather than a true rotation (fig. 4).

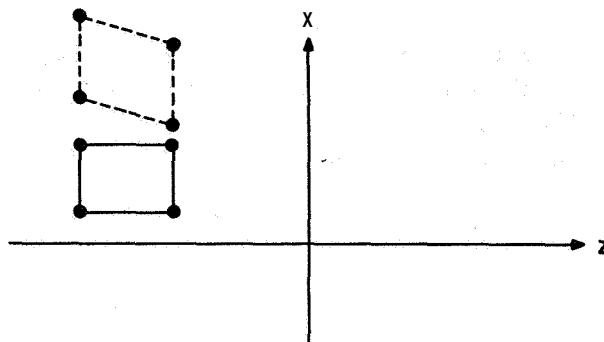


Figure 4.—Shear effect resulting from viewing an object through a thin prism. Each corner of the square images according to the relation $\Delta x = \alpha_z$, $\Delta z = 0$.

HOLOGRAPHIC IMAGE MANIPULATION

In general, we desire six degrees of freedom (three translational and three rotational) for the image motion in order to display the visual effect of arbitrary motion of the observer relative to the scene. In this section we develop the techniques for manipulating the virtual image resulting from a hologram so as to obtain these **six** degrees of freedom.

Recall that the field produced over the aperture of the hologram is given by $E_H = TE_R$. In the previous section we saw that by using the same wave for reconstructing the hologram as was used in construction, we were able to reproduce the original object field. Suppose, however, that the reconstructing and constructing waves are related by the expression $E_R = \gamma(x, y)E_0$. Then, upon reconstruction, the field over the hologram aperture is⁵

$$E_H = \gamma(x, y)E \quad (6)$$

Thus, the observer looking at the hologram no longer sees an exact image of the original object; he sees, rather, an image which is modified in accordance with the function γ . Clearly, this function is analogous to the transmission function discussed previously. Indeed, the effect of using a reconstructing wave which differs from the constructing wave is equivalent to viewing the actual object through an optical system with transmission function 4.

⁵We have assumed that the zero-order and first-order real images are sufficiently separated so that we may neglect their effects.

If γ is allowed to vary with time (i.e., we use a time-varying reconstruction wave), then it is possible to obtain a dynamic reconstructed image. The result is analogous to looking at the object through an optical system with time-varying properties; although the object is fixed, the image which is seen by the observer changes according to the optical properties at the instant of observation. It is important to note that γ need not depend explicitly on time but can be, instead, a function of variables which are time-varying. The approach which could be used in a flight simulator, for example, is illustrated in figure 5. In this case, γ is a function of the aircraft's attitude and position relative to the scene (the Euler angles, φ, θ, ϕ , and the Cartesian position coordinates, x, y, z). Hence, in order to obtain the desired dynamic display, our problem reduces to that of selecting the appropriate function γ . To this end, let us examine the image motion resulting from various simple, realizable selections for γ . (Our technique is, actually, to select various realizable reconstructing waves and determine the corresponding γ .)

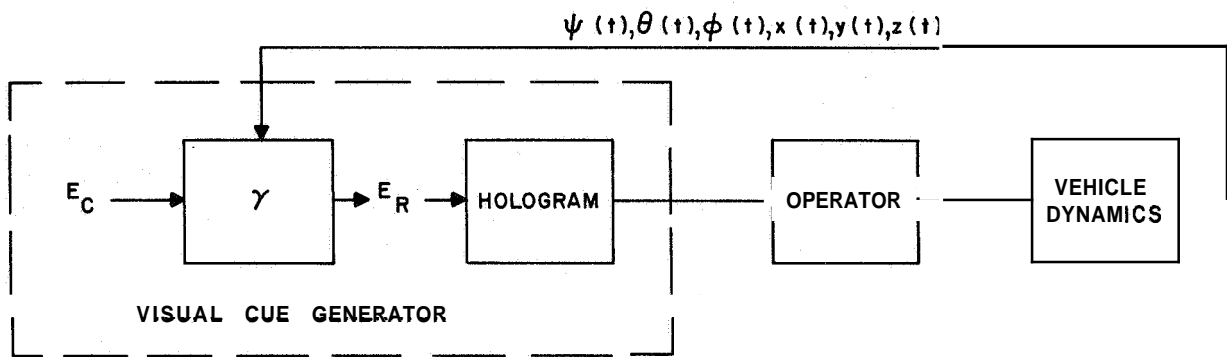


Figure 5.—Simplified typical holographic flight simulator configuration.

Assume that the hologram has been constructed using a unit magnitude plane wave, that is,

$$E_c = \exp j \frac{2\pi}{\lambda} (\alpha_1 x + \alpha_2 y)$$

where α_1 and α_2 are the direction cosines of the normal to the plane wave. This assumption will be used throughout the remainder of the paper. Then

$$\gamma(x, y) = E_R(x, y) \exp -j \frac{2\pi}{\lambda} (\alpha_1 x + \alpha_2 y) \quad (7)$$

Let us first consider the system shown in figure 6 and assume that $E_R' = E_c$. The effect of the aperture is to restrict the area of illumination of the hologram, so that

$$E_R = E_c \quad \forall (x, y) \in R_0$$

$$E_R = 0 \quad \forall (x, y) \notin R_0$$

where R_0 is the aperture region. For this important but trivial case, we obtain

$$\gamma(x, y) = 1 \quad \forall (x, y) \in R_0$$

$$\gamma(x, y) = 0 \quad \forall (x, y) \notin R_0$$

This choice of γ and E_R results in an image which corresponds to viewing the object through a "window" coincident with R_0 . By translating the hologram parallel to the aperture one obtains the same effect that would result from moving the window, relative to the scene, in the plane over which the hologram was constructed.

Moreover, it is apparent that by rotating the hologram about any normal to its surface, we obtain an effective rotation (roll) of the scene about that normal. Thus, by introducing an aperture and by moving the hologram relative to this aperture, we are able to obtain three (one longitudinal, two lateral) of the required six degrees of freedom.

Let us now consider the effect of varying the direction of propagation of the reconstructing plane wave. A possible system for implementing this process is shown in figure 7. Then

$$E_R = \exp j \frac{2\pi}{\lambda} (\alpha'_1 x + \alpha'_2 y)$$

and, from equation (7),

$$\gamma = \exp j \frac{2\pi}{\lambda} [(\alpha'_1 - \alpha_1) x + (\alpha'_2 - \alpha_2) y] \quad (8)$$

But, referring to equation (3), we see that γ in this case is simply the transmission function for a pair of thin prisms. Thus, the effect of changing the direction of the reconstructing plane wave is a "rotation" of the image through angles $(\alpha'_1 - \alpha_1)$ and $(\alpha'_2 - \alpha_2)$. As mentioned in the section on optics, a true rotation is not obtained and the technique is only useful for reasonably small angular rotations. In such cases, however, this method may be used to obtain two additional degrees of rotational freedom (yaw and pitch).

Consider now the case where the reconstructing wave is produced by a point source located at coordinates (x_0, y_0, z_0) (fig. 8). For this case, we obtain, after algebraic manipulation:

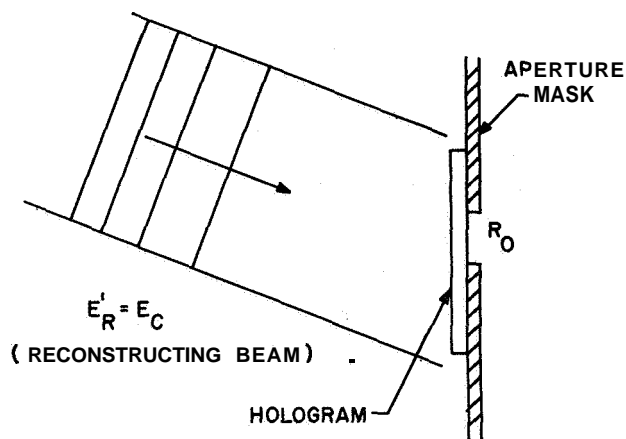


Figure 6.—Technique of reducing the region of illumination.

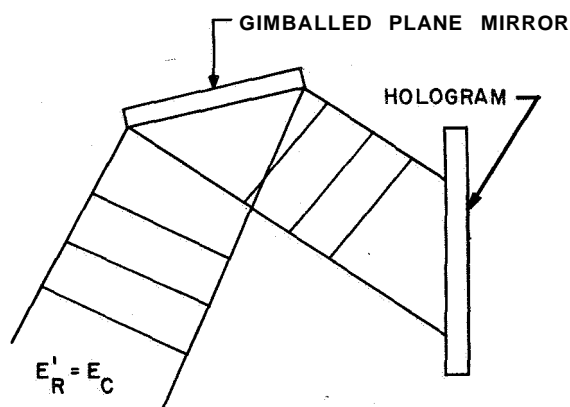


Figure 7.—Changing the angle of a reconstructing plane wave to obtain image rotation.

$$\gamma(x,y) = \exp j \frac{2\pi}{\lambda} \left[\left(\frac{x_0}{z_0} - \alpha_1 \right) x + \left(\frac{y_0}{z_0} - \alpha_2 \right) y - \frac{(x^2 + y^2)}{2z_0} \right]$$

Referring to equations (3) and (4), we see that γ is the transmission function of a thin lens-thin prism system. The equivalent lens has a focal length of z_0 and the prisms have prism angles of $\sin^{-1} \left(\frac{x_0}{z_0} - \alpha_1 \right)$ and $\sin^{-1} \left(\frac{y_0}{z_0} - \alpha_2 \right)$.

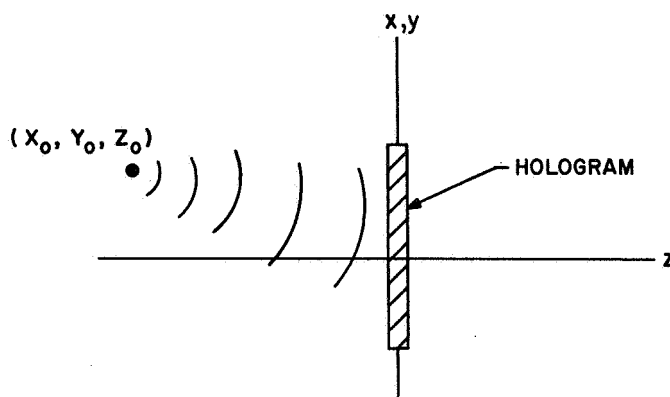


Figure 8.—Using a point source as the reconstructing wave.

Thus, by changing the position of the point source, it is possible to obtain two rotational degrees of freedom, demagnification, and range change (i.e., motion along a perpendicular to the plane over which the hologram was constructed.) Again, it must be stressed that the image motion that results from this technique is not free from distortion. For a given application, the equations (5) must be employed to obtain quantitative information on the amount of distortion actually introduced. It would appear that this last technique would provide the required sixth degree of freedom (i.e., motion along the per-

pendicular to the plane over which the hologram was constructed). However, the net effect of the concurrent demagnification and range change is not, in general, equivalent to the desired degree of freedom. The reason for this phenomenon is discussed in the next section.

We have now seen how image manipulation is possible by varying the reconstructing beam. It is also possible to obtain image motion by placing an optical system between the observer and the hologram as illustrated in figure 9. The optics might, for example, consist of a gimbaled mirror to obtain true rotations or a zoom lens to obtain true range variation.

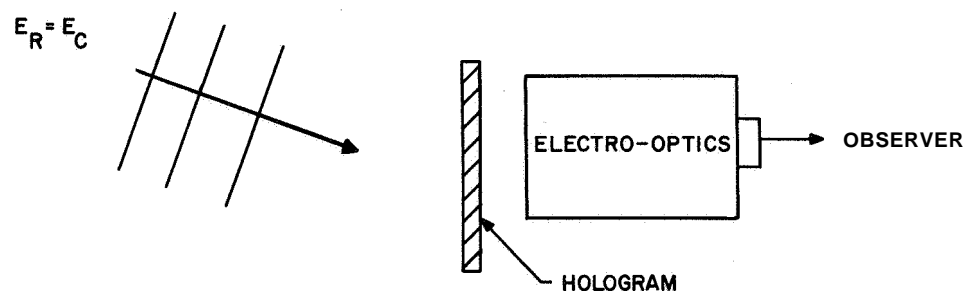


Figure 9.—Image manipulation via optics placed between hologram image and observer.

These techniques can overcome some of the difficulties discussed above, but since they are well known display techniques, we will not treat them here. It is well to point out however, that although these techniques sometimes introduce less distortion, they frequently have the

disadvantages of increased size, complexity, and cost. Of course, both types of image manipulation techniques can be used in combination.

CURRENT LIMITATIONS AND PROBLEMS

The major problem which limits the use of holograms in display applications relates to the maximum object size which can be hologrammed. A hologram is an interference pattern and as such requires the light source to be coherent in both time and space. The degree of coherency (e.g., the spectral purity of the laser output) affects the contrast of the interference pattern and consequently the quality of the reconstructed image. A perfectly coherent source can produce perfect holograms, while a completely noncoherent source (e.g., a white light bulb) can not produce a hologram. The actual sources that are used in holography lie in between these **two** extremes and are referred to as "partially coherent sources." For our purpose, we can consider partially coherent sources to be point sources which emit light over a small band of wavelengths.

Now, consider the holographic construction process illustrated in figure 10. Rays ABD, and ACD are reflected by the mirror and object, respectively, and interfere at point D.

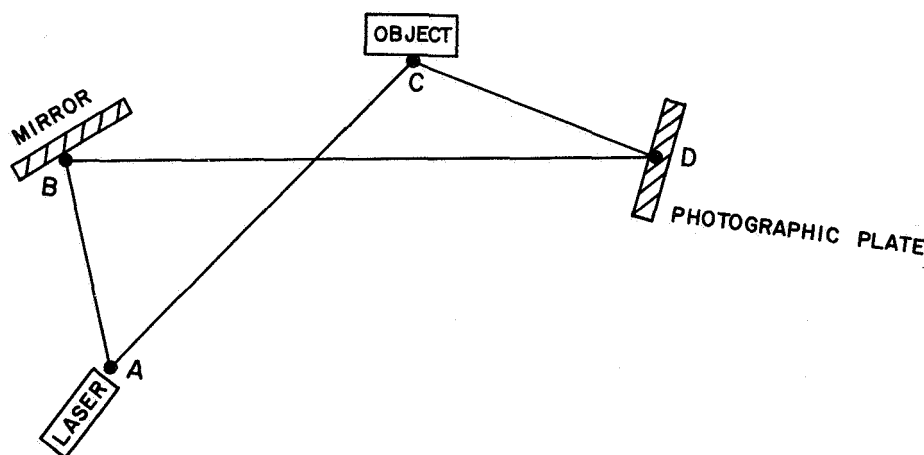


Figure 10.—Typical configuration used in constructing a hologram.

To produce a successful hologram of the object requires that interference does indeed take place at point D thereby allowing the phase of the scattered ray CD to be recorded. It can easily be shown that, for a laser of spectral width $\Delta\lambda$ and center wavelength λ , the lengths L_1 and L_2 of paths ABD, and ACD, respectively, must satisfy the following inequality (ref. 10):

$$|L_1 - L_2| \leq \frac{\lambda^2}{\Delta\lambda}$$

The quantity $\lambda^2/\Delta\lambda$ is referred to as the coherence length of the laser. This length places an upper limit on the maximum object and maximum hologram plate sizes that can be used.

To obtain interference at a given point not only requires a source with sufficient coherence length, but also requires extreme mechanical stability between the object, hologram plate, and laser. Any random vibration between these components will generate a random phase factor between any two arbitrary rays. The result of this random phase is, of course,

a decrease in coherence and consequently a decrease in the degree of interference. Many experimentalists believe this stability problem to be the major problem limiting the quality of realizable holograms (ref. 11).

The end result of the above problems is to limit the object size and hologram plate size. The effect of these size limitations on displays, such as those for flight simulators, is that one is limited to using holograms of small models, or small sections of large models, in the simulator. This limitation is more severe than one would at first suspect. In particular, it limits the applicability of holograms to three dimensional simulators. The reason for this can be understood by referring to figure 11. In this figure we show two congruent objects located at distances L_1 and L_2 relative to a viewing window. As an example, suppose O_2 is an

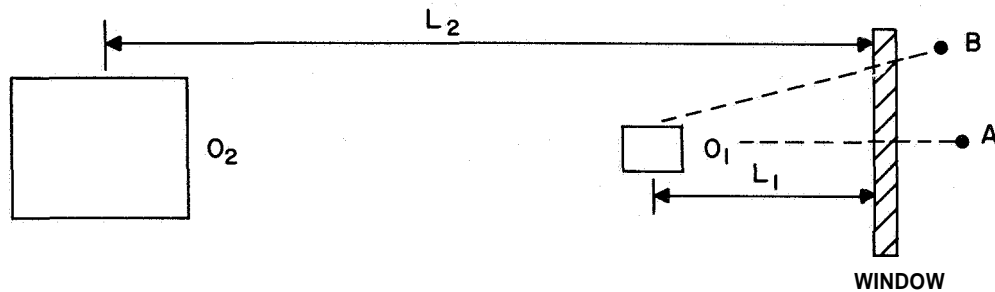


Figure 11.—Result of using a scale model O_1 of an object O_2 in a display system.

airport and the window is the windshield of an aircraft. Let O_1 be a reconstructed image of a hologram made from a small scale model of the airport and the window be the hologram itself. By properly choosing L_1 and L_2 , it is apparent that the image size formed on the retina of the eye can be made identical in the two cases. Thus, if we wished to simulate landing at an airport, we could use a television camera looking into the hologram and view the image on a television monitor. Except for model imperfections, etc., one could not distinguish whether he were viewing a television picture of the hologram image or of the actual airport. If, however, we looked directly into the hologram and moved our eyes from point A to point B, we would be able to "see around" the hologram image, whereas we could not see around the actual object. A solution would be to scale **both** the model size **and** the hologram window size. Simulators, however, usually require the viewing window to be the same size as that of the vehicle which is being simulated. Consequently, the models will be required to have the exact dimensions of the actual objects, or the actual objects themselves will have to be hologrammed. To accomplish this obviously requires large coherence lengths and extreme stability. Further progress in the "art" of constructing holograms will certainly be required before this becomes feasible.

It was mentioned in the previous section that range change is, in general, not possible with the techniques which have been presented. If, however, a two-dimensional display is required (e.g., television picture of the hologram image), one can use the spherical reconstructing wave to effect this sixth degree of freedom. Thus, although the image moves closer to the observer, it also decreases in size. The size decrease, when viewed on a television monitor, will appear to be equivalent to a range increase. Since one is largely restricted to two-dimensional holographic simulations because of the problems discussed above, it is seen that in terms of presently realizable holographic simulators, full six-degree-of-freedom operation is possible.

CONCLUSION

We have discussed basic concepts in holography and optics and have seen that holographic systems have properties which make them quite attractive for flight displays and flight simulators. In order to obtain a dynamic display from a single hologram, we investigated techniques for manipulating the virtual image resulting from a hologram. It was shown that it is possible to obtain a reasonably large amount of image motion by simple techniques such as moving the hologram relative to a fixed aperture, or adjusting a spherical reconstructing wave. The resulting motion is equivalent to observer motion over limited regions of space. The bounds on the region would depend on factors such as the amount of distortion which can be tolerated, the size of the hologram, the size of the aperture, and the characteristics of the optics used in forming the reconstructing beam.

Some of the problems associated with holographic display systems were also discussed. The principal problem was seen to be the limitation imposed on model and hologram size. This limitation is due to the relatively small coherence lengths of present lasers and the stability requirements for holographic construction. A method for overcoming the resultant display difficulty involves the use of a closed-circuit television system. This solution however, is not entirely satisfactory since the technique involves a loss of some of the more desirable display features inherent in holograms. Alternative solutions to this problem must be found if the full potential of holographic display systems is to be realized.

Finally, it is important to remember that, although the principles of holography have been known for some time, holographic technology is still in its infancy. Rapid strides have been made recently and we can expect more of the same. It seems entirely plausible that, in the not too distant future, it will be possible to display an image, using holographic techniques, which is indistinguishable from the real-world scene.

REFERENCES

1. Gabor, D.: Wavefront Reconstruction or "Holography." Proc. Eighth Annual Electron and Laser Beam Symposium, Apr. 6, 1966.
2. Gabor, D.: Microscopy by Reconstructed Wave Fronts. Proc. Roy. Soc., A197, 1949.
3. Stroke, G. W.; and Labeyrie, A. E.: White Light Reconstruction of Holographic Images Using the Lippman-Bragg Diffraction Effect. Phys. Rev. Letters, vol. 20, no. 4, Mar. 1, 1966, p. 368.
4. Wilmot, D. W.; Schineller, E. R.; and Heuman, R. W.: Hologram Illumination With a Flashlight. IEEE Proc., vol. 54, Apr. 1966, pp. 690-691.
5. Lin, L. H.; Pennington, R. S.; Stroke, G. W.; and Labeyrie, A. E.: Multicolor Holographic Image Reconstruction With White Light Illumination. Bell Systems Tech. J., Apr. 1966, p. 659.
6. Born, M.; and Wolf, E.: Principles of Optics. Second rev. ed., Pergamon Press, 1964.
7. Stroke, G. W.: Introduction to Coherent Optics and Holography. Academic Press, 1966.
8. Leith, E. N.; and Upatnieks, J.: Wavefront Reconstruction With Continuous-Tone Objects. J. Opt. Soc. Am., vol. 53, 1963, pp. 1377-1381.
9. Meier, R. W.: Aberrations and Symmetry Properties of Holographic Images. J. Opt. Soc. Am., vol. 59, 1965, p. 1566.
10. Fleisher, H.; Pengelly, P.; et al.: Optical and Electro-Optical Information Processing. Ch. 1, MIT Press, 1965.
11. Morrow, H. E.; and Dessel, N. F.: Three Dimensional Holography. NEL 1403, U. S. Navy Electronics Lab., Sept. 7, 1966.

5. Human Operator Describing Functions with Visual and Tactile Displays*

James C. Bliss
Stanford Research Institute

768 15906

In the attempt to develop models of manual tracking behavior that also incorporate characteristics of the physiological systems underlying the responses, techniques are needed to separate the contributions of the individual physiological systems, including sensory, central, or motor functions. Techniques that have been used in the past involve variations in the type of command signal (e.g., transient, periodic, and random), the type of output response (e.g., continuous or discrete), and the vehicle dynamics. The research described in this paper emphasizes varying the sense modality employed (i.e., visual, tactile, or both) with continuous command signals and pure-gain vehicle dynamics.

Two experiments are reported. The first experiment compares the describing functions obtained with (1) a visual display, (2) a tactile display, and (3) both displays used simultaneously. The second experiment explores various tactile display conditions.

PROCEDURES

A system for performing tactile and visual tracking experiments and obtaining amplitude and phase measurements of the response as a function of frequency was developed. This system consists of CDC 8090 programs, A/D and D/A conversion channels, and display and response apparatus.

The CDC 8090 computer programs consist of several parts: a signal generator program, a signal analysis program, and an INTERFOR' program for calculations on the data. The first two parts cyclically generate a value for the D/A output signal and analyze a response value input through the A/D channel. At the end of an adjustable time, the experimental trial is terminated and control is transferred to an INTERFOR program, which performs some calculations on the analysis results and outputs the amplitude and phase values on the on-line typewriter.

The signal generator program contains a table of 15 values representing a quarter cycle of a sine wave. The program uses this table to generate a composite signal, consisting of a sum of sinusoids of arbitrary amplitude and phase. Thus,

$$c(t_k) = \sum_i^N c_i \sin(\omega_i t_k + \phi_i)$$

*This research was sponsored by the National Aeronautics and Space Administration under Contracts NAS 2-2752 and NAS 2-3649.

INTERFOR is a Control Data programming system for the CDC 160-A Computer that is relatively easily interfaced with machine language programs.

where $c(t_k)$ is the value of the generated signal during interval t_k , c_i is the amplitude, φ_i is the phase, and ω_i is the frequency of the i th sinusoid. Up to eight frequencies can be accommodated by the program, and since the longest program cycle is 75 msec, frequencies up to 6.7 cps are practical with the full eight-signal sinusoids plus eight additional analysis sinusoids. (Higher frequencies are practical if fewer than 16 (total) sinusoids are used.)

The accuracy of the A/D and D/A conversion is eight bits. The arbitrary amplitude of each generated frequency component can be selected with fourbits. The eight frequencies can be chosen with fourbits, and the phases can be chosen with four bits within each quarter cycle.

By positioning a sense switch on the computer console, either the internally generated signal can be subtracted from the response and the difference (i.e., the error) outputted (e.g., for compensatory tracking), or the internally generated signal (i.e., the command) can be outputted directly (e.g., for pursuit tracking).

In the analysis programs, an input signal is multiplied by each of a number of sine and cosine components, consisting of the frequencies generated by the signal generator program plus up to eight additional frequencies. Cumulative sums of the results of these multiplications are updated each program cycle. Thus, if the input signal is $r(t_k)$, then the sums a_j and b_j are formed as follows:

$$a_j = \sum_{k=0}^T c_j \sin(\omega_j t_k) r(t_k)$$

$$b_j = \sum_{k=0}^T c_j \cos(\omega_j t_k) r(t_k)$$

The input signal to this program is either the response signal directly from the A/D converter or the difference between the response and the generator signal (i.e., the error), depending on the position of a console switch.

The clock for the programs described above is either internally based on the computer memory-cycle time or controlled by an external pulse generator, depending on the position of a console switch. Shorter program-cycle times are possible with the internal clock, but greater time accuracy is possible with the external pulse generator.

The INTERFOR program takes the sums generated during the experimental trial by the analysis program and computes the amplitude r_j and phase φ_j of each of up to 16 frequency components according to the following equations:

$$r_j = \frac{2}{T} \sqrt{a_j^2 + b_j^2}$$

$$\varphi_j = \tan^{-1} \frac{a_j}{b_j}$$

The results of these calculations are then typed out on the on-line typewriter.

Thus, pursuit or compensatory tracking experiments can be performed with real-time determination of either response or error spectra. Up to eight sinusoids can be used to

generate the command signal, and an additional eight sinusoids can be used to determine the characteristics of the remnant. In addition, the total power in the response is computed so that the correlation between the response and the corresponding linear system can be determined.

A continuous tactile display system was constructed for tactile compensatory tracking. This display consisted of a servo-positioned airjet stimulator which moved horizontally across the forehead or the palmar side of the hand over a range of about 4.5 inches. The airjet was positioned according to the computer-generated signal from the D/A channel. Figure 1 shows this display system when used as a forehead display, and figure 2 shows the display adapted for stimulation on the palmar side of the hand. A visual display that was analogous to the forehead tactile display was obtained by placing a mirror in front of the subject so that he could see the arm that carried the airjet nozzle. A stationary pointer was attached to the forehead rest to give a zero reference for the visual display. The visual counterpart to the tactile display for the palmar side of the hand was obtained by having the subject merely watch the airjet nozzle directly. Again, a pointer was provided to give a visual zero reference.

The computer system was calibrated by connecting the output command signal from the D/A converter to the A/D response channel. The analysis was then performed on a "perfect response"—these results agreed within the expected 8-bit accuracy. As a further check, the system was used to measure the Bode diagrams for a simple resistor and capacitor divider forming a 1-cps low-pass filter. In figure 3, the computer analysis, the measured values from applying each sinusoid individually, and the calculated values are compared.

The computer system was then used to measure the Bode diagrams for the servo system, and these results are shown in figure 4. In all of the subject-describing functions, the servo characteristics were removed either by calculation or by analyzing the feedback pot signal and considering this to be the error input to the subject.

Table 1 gives the frequencies and amplitudes that composed the command signal. Three male subjects in their early twenties were used. Previous to testing, each subject was given from 12 to 30 (depending on his mean squared error scores) 2-minute trials, during which the display was changed in alternate trials from tactile to visual.



Figure 1.—Tactile tracking apparatus for forehead.

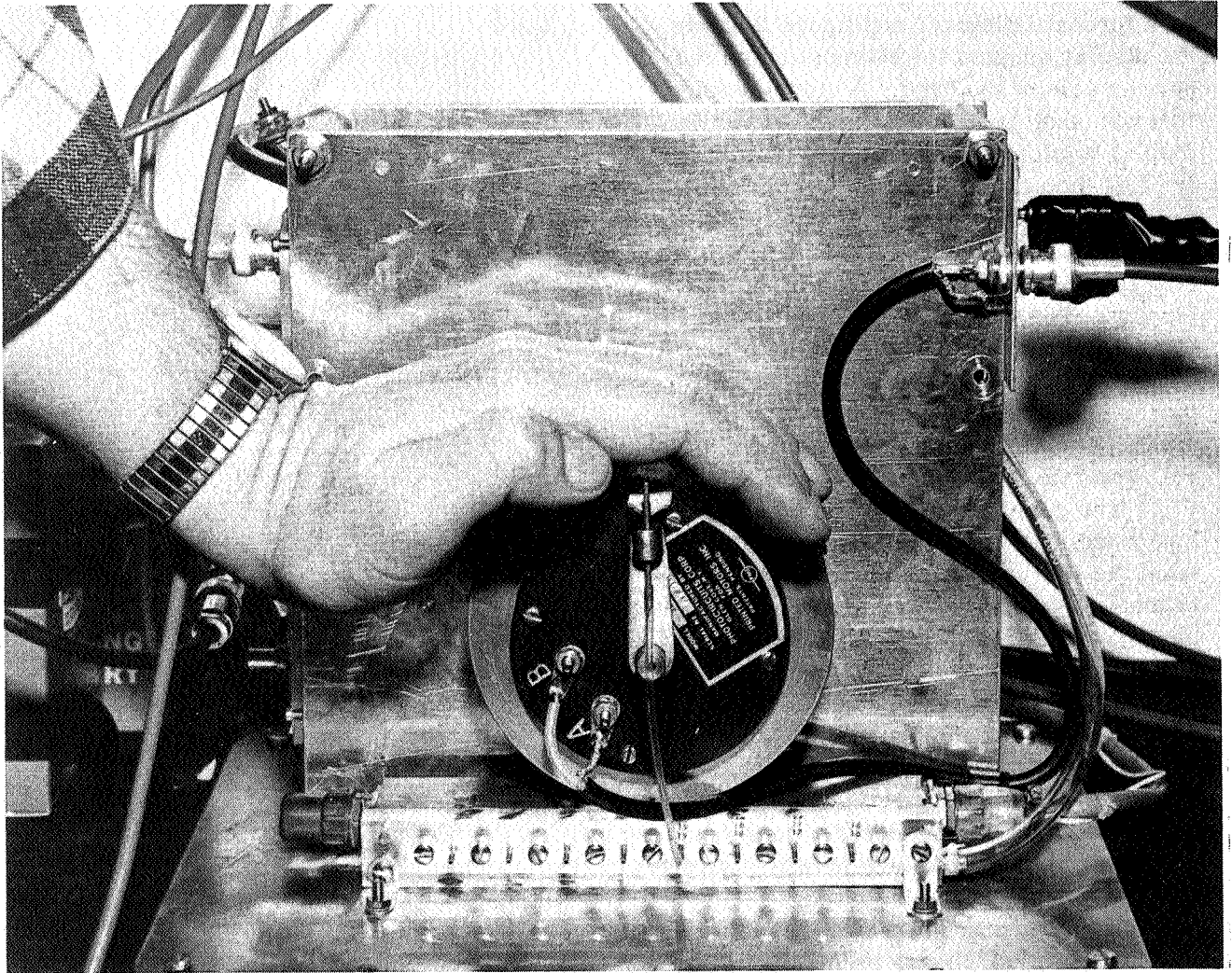


Figure 2.—Continuous tactile tracking display for the palmar side of the hand.

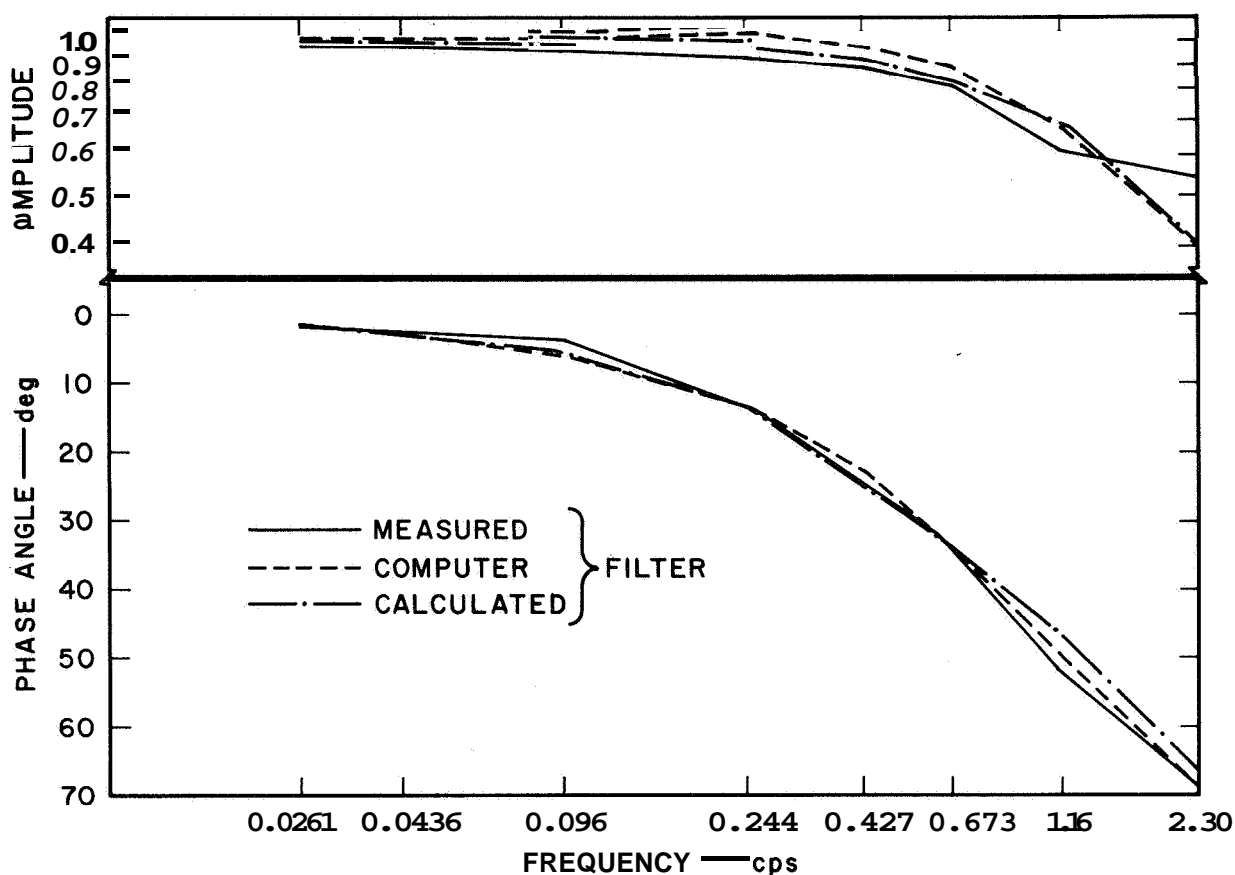


Figure 3.—Comparison of amplitude and phase characteristics versus frequency of a 1-cps filter determined by direct measurement, computer tracking system analysis, and theoretical calculation.

EXPERIMENTS

EXPERIMENT 1—TRACKING WITH VISUAL AND TACTILE DISPLAYS.—In this experiment, forehead airjet tactile tracking was compared with tracking by visually observing the tactile display through a mirror. The airjet stimulator was not activated in the visual tracking runs, but six additional tactile stimulators for auditory masking were used in both the tactile and visual tracking runs.

Two sessions were run with each subject. In the first session, twelve 4-minute tracking runs were performed, alternating between the tactile and visual displays. In the second session, six 4-minute tracking runs were performed, alternating between the tactile, visual, and both display conditions.

The individual subject variability among trials was comparable to the variability across subjects; therefore, only data averaged over all sessions and subjects with the standard deviation at each point are presented in figure 5. While there appears to be no difference between the visual and both-display conditions, tracking with the tactile display resulted in much less low-frequency gain and a reduced crossover frequency.

EXPERIMENT 2.—VARIOUS TACTILE DISPLAY CONDITIONS.—In view of the relatively poor performance with the tactile display in experiment 1, several modifications were made to determine whether the tactile tracking could be improved. A single subject who gave the

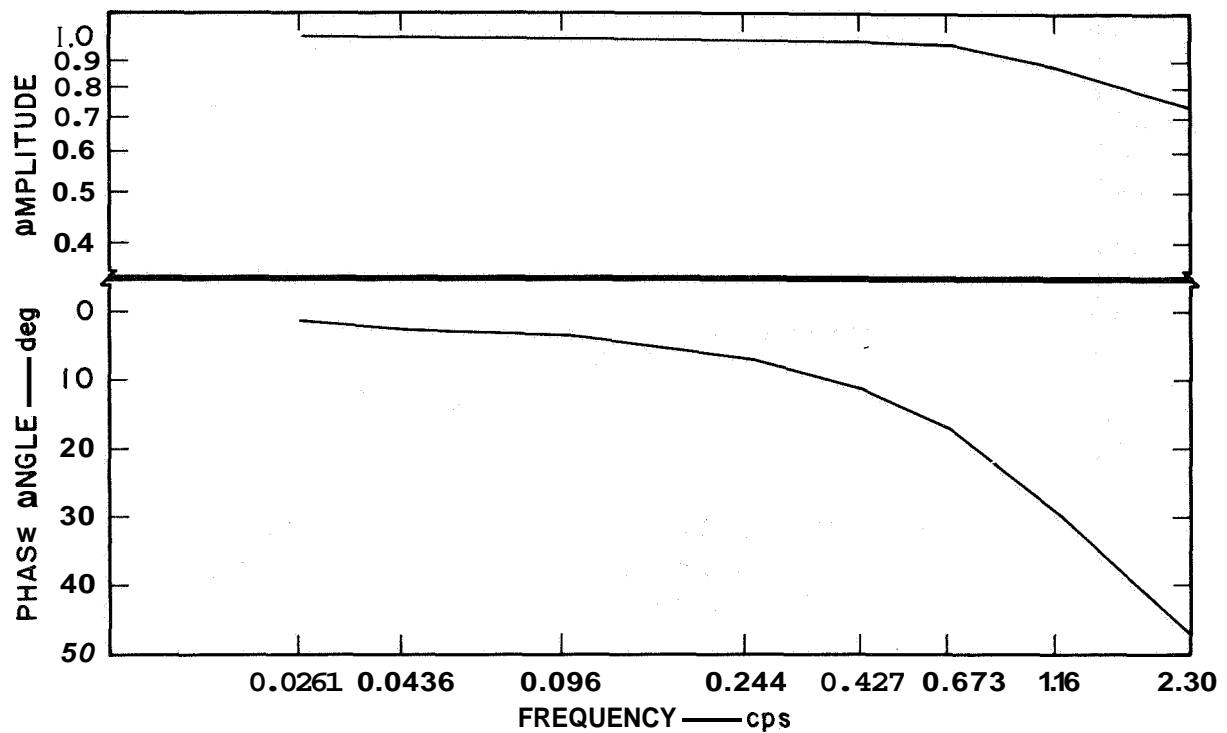


Figure 4.—Bode plots for display servo system as determined by the computer tracking system.

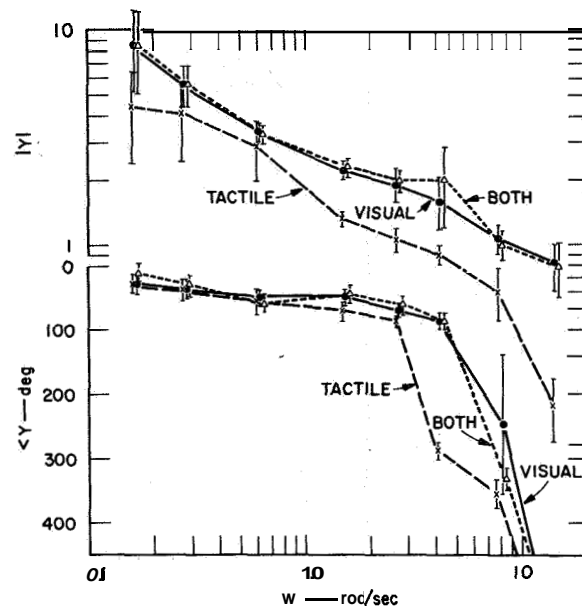
most consistent behavior in the experiment was selected for these explorations. For these sessions, the location of the tactile stimulation was changed to the palmar side of the hand. In addition, the frequency of the airjet stimulation was adjusted to 40 cps and then to 70 cps. Finally, the airjet stimulator was turned off, and the nozzle tip was adjusted so that it touched the palmar side of the hand lightly.

TABLE 1.—COMPOSITION OF COMMAND SIGNAL

Frequency		Amplitude
cps	rad/sec	
0.0261	0.164	1.0
.0436	.274	1.0
.0960	.603	1.0
.2440	1.53	1.0
.4270	2.68	1.0
.6730	4.23	1.0
1.25	7.85	.25
2.30	14.45	.25

Figure 6 shows the describing functions obtained for each of these conditions. There appears to be little difference between the results with the forehead stimulation and the hand

Figure 5.—Describing functions, averaged over subjects and sessions, for three display conditions.



stimulation, and among the results with the various frequencies of airjet pulsation; however, significant improvement was obtained with the contact stimulus,

RESULTS AND DISCUSSION

COMPARISON OF FOREHEAD AND HAND TACTILE DISPLAYS.—Figure 7 shows the forehead and hand tactile results averaged over the three subjects. The differences between these two curves are less than one standard deviation, except near the crossover frequency, where the difference is slightly more than one standard deviation. Thus these differences are hardly significant, and it appears that forehead stimulation and hand stimulation resulted in approximately equal performance.

COMPARISON BETWEEN VISUAL AND TACTILE PERFORMANCE.—The best performance with a tactile display was obtained on the hand when the airjet nozzle was turned off, but allowed to contact the skin. Figure 8 shows a comparison between this tactile-contact condition and visual performance for one subject. The amplitude differences are significant, but the phase curves are practically identical.¹ Thus, our tentative results, based on one subject, suggest that with tangential as well as normal forces on the skin, the tactile performance has equal bandwidth, but less gain, than the visual performance.

MEAN-SQUARE ERROR AND DISPLAY MEASUREMENTS.—The computer also calculated the mean-square error and display for each run. Tables 2 and 3 show these results, averaged over three sessions, for each subject and each condition of the experiment. Tactile mean-square error was generally about twice the visual mean-square error. The one anomalous result was with subject 3, who produced a tactile (hand display) mean-square error of more than five times his visual mean-square error.

¹It is interesting to compare figure 8 with figure 2 of paper 2 in this volume.

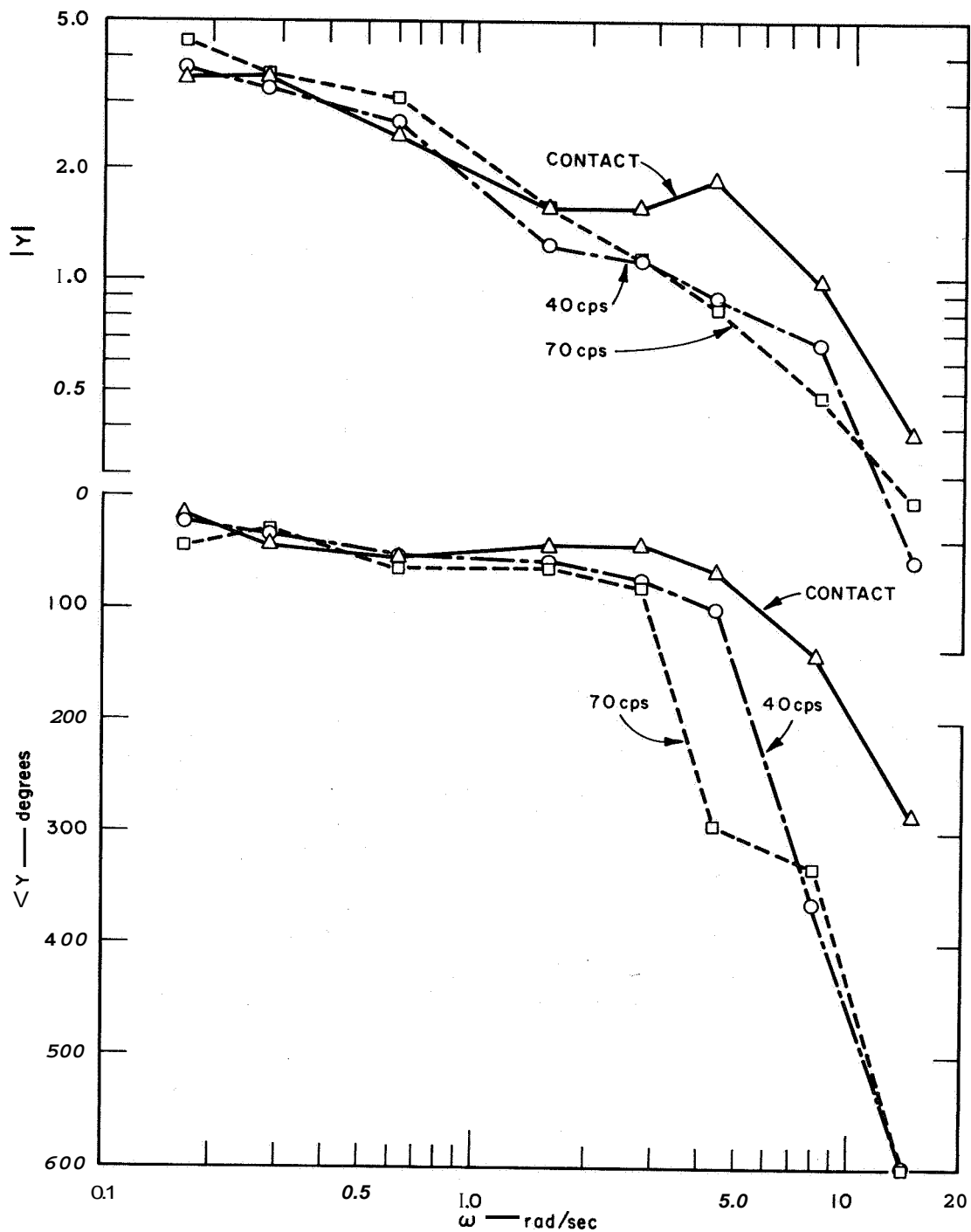


Figure 6.—Describing functions for three conditions of the tactile display.

Although the describing function results indicate superior performance under the tactile-contact condition, the mean-square error values for the 200 cps and 70 cps conditions were lower.

In general there was less mean-square error when both displays were used simultaneously than with either display alone.

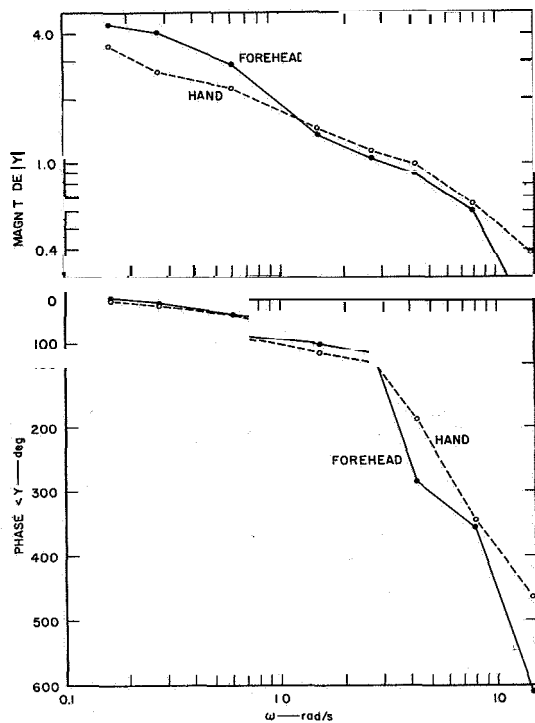


Figure 7.—Describing functions for hand and forehead tactile displays (200 cps airjet).

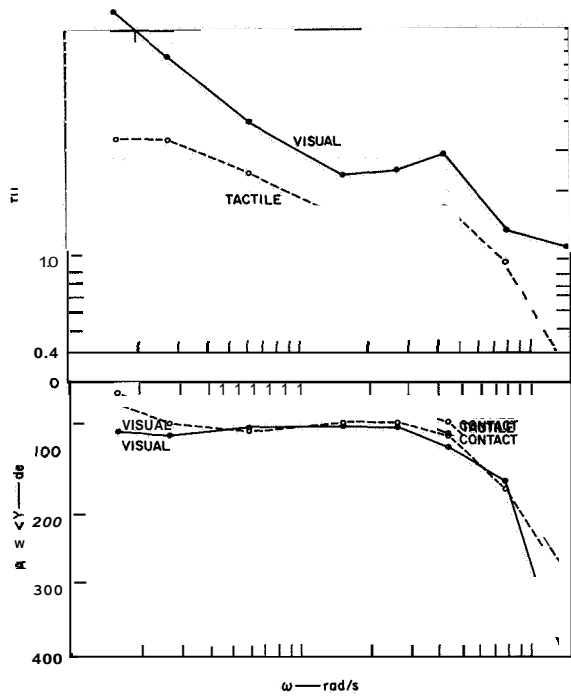


Figure 8.—Describing functions for visual and tactile (hand-contacting stimulus) displays.

Subject	Visual	Tactile	Both
1	2.32	4.91	1.84
2	2.06	5.69	2.10
3	2.63	5.88	2.40
Average	2.33	5.49	2.11

Subject	Visual	Tactile			
		200 cps	70 cps	40 cps	Contact
1	1.58	3.04	3.06	5.39	3.80
2	1.63	4.04	--	--	--
3	1.94	11.3	--	--	--
Average	1.72	6.13	--	--	--

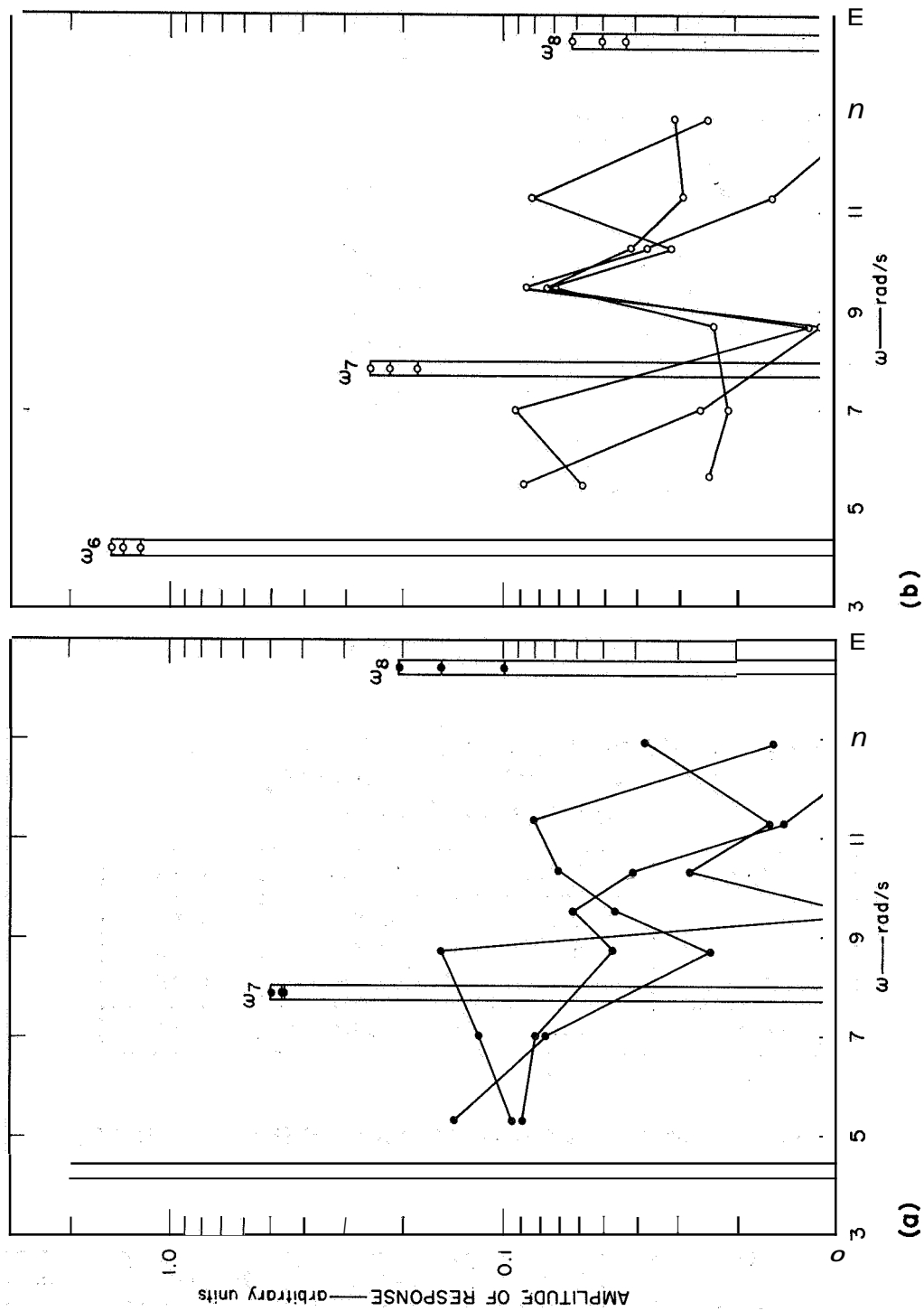


Figure 9.—Amplitude of response as a function of frequency for subject 1 during three runs. (a) Visual display, (b) Tactile display—forehead, 200 cps.

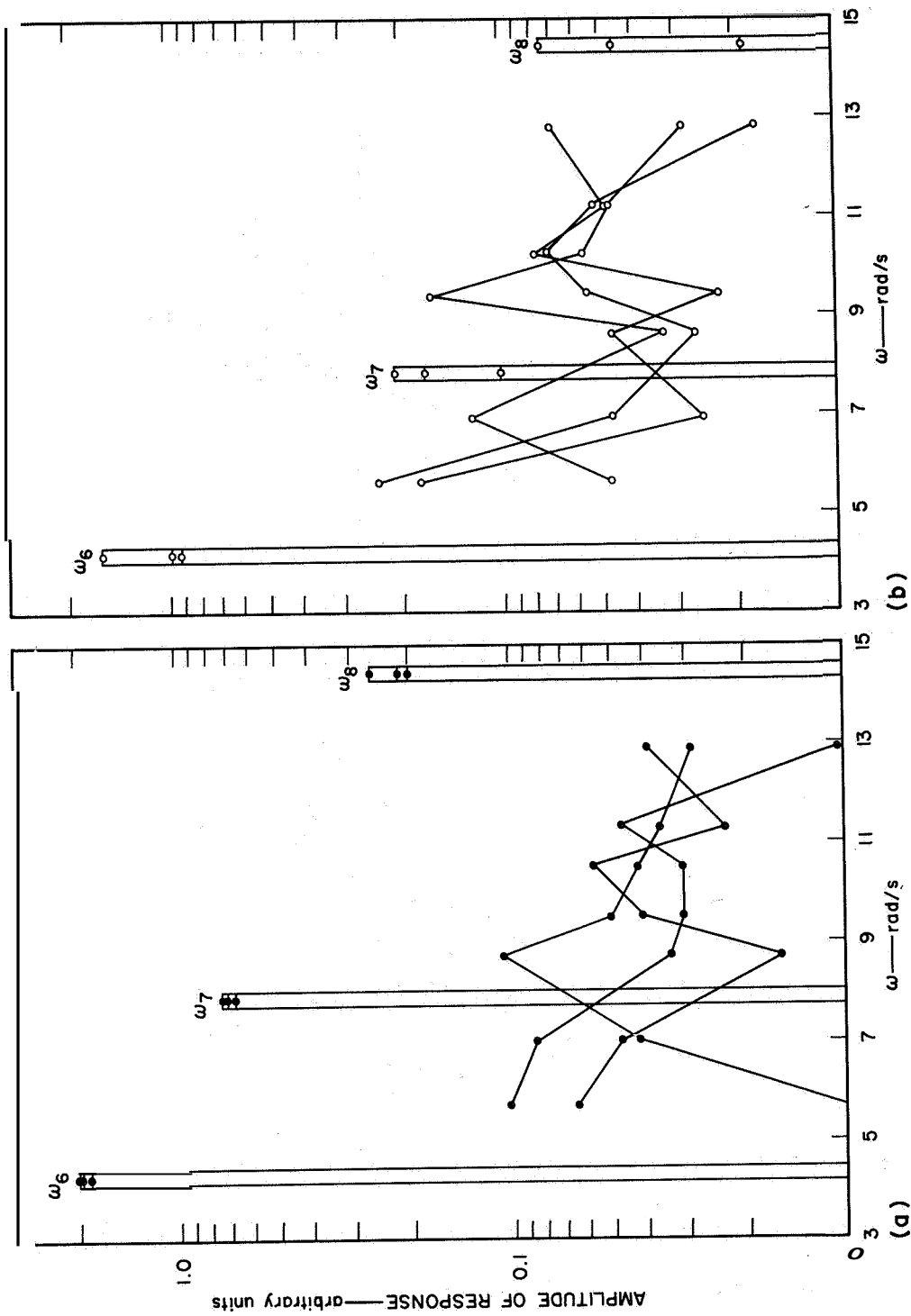


Figure 10.—Amplitude of response as a function of frequency for subject 2 during three runs. (a) Visual display, (b) Tactile display—forehead, 200 cps.

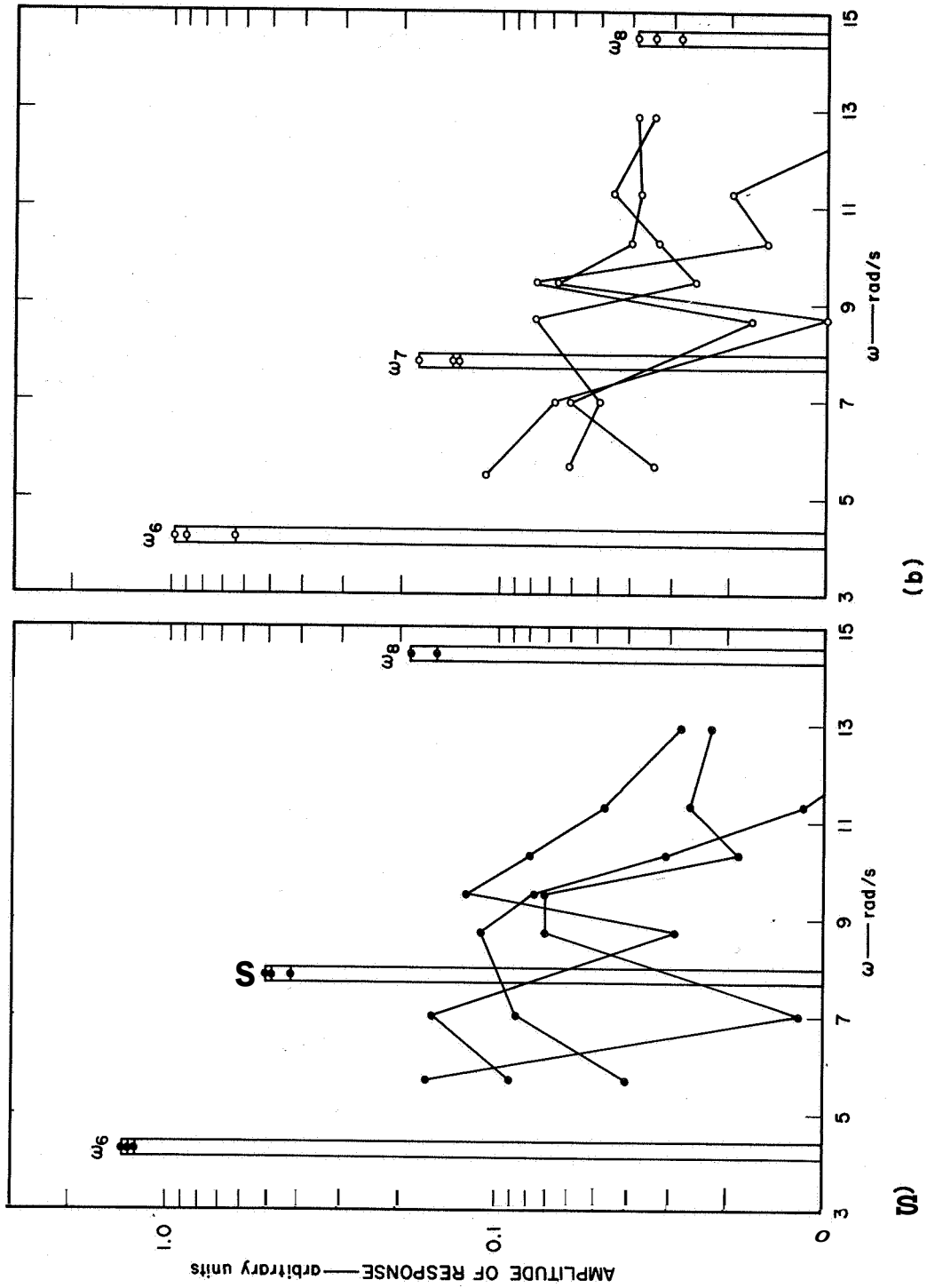


Figure 11—Amplitude of response as a function of frequency for subject 3 during three runs. (a) Visual display, (b) Tactile display—forehead, 200 cps.

REMNANT DATA.—There have been a number of suggestions that a model for the human operator should include a nearly periodic sampler (e.g., ref. 1). The following quotation from McRuer et al. (ref. 2) explains the effect of this hypothesis on the output spectrum:

With the line spectrum forcing function, nonlinearities in the operator would be expected to result in output spectrum peaks which are harmonically related to the forcing function frequencies. Constant-rate sampling on the part of the operator will also tend to produce recurring peaks and valleys in the output spectrum. If the sampler is precisely periodic at a frequency ω_s , output spectral lines would be expected at frequencies $\omega_n \pm m\omega_s$, $m=0, 1, 2, 3 \dots$. Slight variations in sampling rate over a measurement run would tend to slur the lines into peaks.

To examine our data with regard to this hypothesis, the output spectra from single runs for each subject were plotted as shown in figures 9 to 11. The forcing-function frequencies are shown as line spectra. The other measurement frequencies were chosen in the region around half the expected sampling frequency, to coincide with Bekey's "sampling peak" (ref. 1). For this reason, only the region between 4 and 15 rad/s is plotted. The output spectra at ω_7 and ω_8 are uncorrected for the fact that the forcing-function amplitude at these frequencies was only one-fourth that at the lower frequencies.

These data fail to show any stable, clearly defined peaks that could be attributed to a periodic sampling nonlinearity. The power at remnant frequencies is generally less than that at the nearby forcing-function frequencies. Moreover, the remnant curves are certainly not reproducible from run to run.

While the response power at ω_6 , ω_7 , and ω_8 with the visual display is consistently greater than that with the tactile display, the remnant power is roughly the same with the two displays.

REFERENCES

1. Bekey, G. A.: The Human Operator as a Sampled-Data System. IRE Trans., HFE-3, Sept. 1962, pp.43-51.
2. McRuer, D.; Graham, D.; Krendel, E.; and Reisener, W., Jr.: Human Pilot Dynamics in Compensatory Systems—Theory, Models, and Experiments With Controlled Element and Forcing Function Variations. AFFDL-TR-65-15, July 1965.

2 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82 85 88 91 94 97 100 103 106 109 112 115 118 121 124 127 130 133 136 139 142 145 148 151 154 157 160 163 166 169 172 175 178 181 184 187 190 193 196 199 202 205 208 211 214 217 220 223 226 229 232 235 238 241 244 247 250 253 256 259 262 265 268 271 274 277 280 283 286 289 292 295 298 301 304 307 310 313 316 319 322 325 328 331 334 337 340 343 346 349 352 355 358 361 364 367 370 373 376 379 382 385 388 391 394 397 400 403 406 409 412 415 418 421 424 427 430 433 436 439 442 445 448 451 454 457 460 463 466 469 472 475 478 481 484 487 490 493 496 499 502 505 508 511 514 517 520 523 526 529 532 535 538 541 544 547 550 553 556 559 562 565 568 571 574 577 580 583 586 589 592 595 598 601 604 607 610 613 616 619 622 625 628 631 634 637 640 643 646 649 652 655 658 661 664 667 670 673 676 679 682 685 688 691 694 697 700 703 706 709 712 715 718 721 724 727 730 733 736 739 742 745 748 751 754 757 760 763 766 769 772 775 778 781 784 787 790 793 796 799 802 805 808 811 814 817 820 823 826 829 832 835 838 841 844 847 850 853 856 859 862 865 868 871 874 877 880 883 886 889 892 895 898 901 904 907 910 913 916 919 922 925 928 931 934 937 940 943 946 949 952 955 958 961 964 967 970 973 976 979 982 985 988 991 994 997 1000

1001 1004 1007 1010 1013 1016 1019 1022 1025 1028 1031 1034 1037 1040 1043 1046 1049 1052 1055 1058 1061 1064 1067 1070 1073 1076 1079 1082 1085 1088 1091 1094 1097 1100 1103 1106 1109 1112 1115 1118 1121 1124 1127 1130 1133 1136 1139 1142 1145 1148 1151 1154 1157 1160 1163 1166 1169 1172 1175 1178 1181 1184 1187 1190 1193 1196 1199 1202 1205 1208 1211 1214 1217 1220 1223 1226 1229 1232 1235 1238 1241 1244 1247 1250 1253 1256 1259 1262 1265 1268 1271 1274 1277 1280 1283 1286 1289 1292 1295 1298 1301 1304 1307 1310 1313 1316 1319 1322 1325 1328 1331 1334 1337 1340 1343 1346 1349 1352 1355 1358 1361 1364 1367 1370 1373 1376 1379 1382 1385 1388 1391 1394 1397 1400 1403 1406 1409 1412 1415 1418 1421 1424 1427 1430 1433 1436 1439 1442 1445 1448 1451 1454 1457 1460 1463 1466 1469 1472 1475 1478 1481 1484 1487 1490 1493 1496 1499 1502 1505 1508 1511 1514 1517 1520 1523 1526 1529 1532 1535 1538 1541 1544 1547 1550 1553 1556 1559 1562 1565 1568 1571 1574 1577 1580 1583 1586 1589 1592 1595 1598 1601 1604 1607 1610 1613 1616 1619 1622 1625 1628 1631 1634 1637 1640 1643 1646 1649 1652 1655 1658 1661 1664 1667 1670 1673 1676 1679 1682 1685 1688 1691 1694 1697 1700 1703 1706 1709 1712 1715 1718 1721 1724 1727 1730 1733 1736 1739 1742 1745 1748 1751 1754 1757 1760 1763 1766 1769 1772 1775 1778 1781 1784 1787 1790 1793 1796 1799 1802 1805 1808 1811 1814 1817 1820 1823 1826 1829 1832 1835 1838 1841 1844 1847 1850 1853 1856 1859 1862 1865 1868 1871 1874 1877 1880 1883 1886 1889 1892 1895 1898 1901 1904 1907 1910 1913 1916 1919 1922 1925 1928 1931 1934 1937 1940 1943 1946 1949 1952 1955 1958 1961 1964 1967 1970 1973 1976 1979 1982 1985 1988 1991 1994 1997 2000

6. A Synthetic Display Technique for Computer-Controlled Simulator and Airborne Displays

Jack J. Hatfield
Langley Research Center, NASA

Interest in airborne, computer-driven, integrated displays has resulted from the problem of effectively displaying greater quantities of rapidly changing control information. A new concept for implementing computer-controlled simulator and airborne displays is described. This concept employs the synthetic generation of desired instrumentation at the cockpit interface of the man-machine loop utilizing a programmable electronic display system. The synthetic display concept utilized is based on an "electronic animation" technique which allows the cockpit display designer to proceed directly from static (cardboard) instrument mockups to dynamic displays which are simulated in the cockpit by high-resolution closed-circuit monochrome television.

Experimental synthetic display hardware is described from the viewpoint of relating the operational capability and flexibility of the "electronic animation" technique. The performance achieved to date is illustrated through photographs of synthetically generated electro-mechanical instrumentation. The potential of the technique for synthesizing experimental displays at lower costs, at higher speed, and in new integrated formats is discussed.

Information display for the manual control of aerospace vehicles is a field which is receiving much attention because of the increasing complexity of pilot's tasks during advanced missions. In complex, multiphased missions, man is faced with greater quantities of control information, events which occur more rapidly, and the requirement for more exacting control. This recognized trend toward increasingly more complex pilot's tasks along with studies which reveal a great deal of pilot scanning activity using conventional instrument arrays (ref. 1) have led to many proposals for an integrated, time-shared display using a computer-driven, general purpose device such as a cathode ray tube (refs. 2 to 4).

A general-purpose display concept seems to offer the potential for more effective display as well as for less equipment and panel space, which would afford lower total weight and volume. In addition, a general-purpose display device is more easily integrated with a central avionics system computer than are a host of individual panel instruments. The capabilities attendant to most computer-generated displays for programing, panel space time-sharing, and display integration could be made available in the cockpit. Yet there is general disagreement among flight control-display system designers as to whether the general-purpose display concept or the conventional instrument array concept offers the greatest potential (ref. 5).

Such disagreements point out the need for sound simulator research programs to determine the effectiveness of general-purpose display concepts and to explore the problems (both from a human factor and hardware standpoint) which would be encountered in the transition from conventional instrument arrays to a small number of central, computer-driven,

general-purpose displays. This paper reports on the investigation and development of a synthetic display technique having applications for programmed, integrated display in simulators and in airborne vehicles. The objective of the work described has been to provide for a more effective, less costly, and less time-consuming means of creating dynamic instrument replicas for use in the simulator evaluation of integrated and general-purpose display concepts.

MOTIVATION FOR SYNTHETIC DISPLAY DEVELOPMENT

THE DISPLAY PROBLEM.—Although the time-proven research simulator is in widespread use both for human factors research and for control-display interface design, it has inherent deficiencies which can be improved upon. The procedure of constructing or procuring physical instruments for use as dynamic replicas has led to problems with development time lags and costs and has produced generally single-purpose instruments. The creation of dynamic instrument replicas can call for many of the following time-consuming techniques (refs. 6 and 7):

- (1) The selection and/or design of many types of electromechanical instruments
- (2) Modification of meter faces or tapes
- (3) Construction and/or modification of servo-driven gear trains
- (4) Construction of special oscilloscope drive circuitry
- (5) Modification of image projectors

In the case of more sophisticated electronic displays, new instrument designs require even longer lead time and can prove to be quite costly. An additional disadvantage to the above-mentioned techniques for implementing research simulator display panels is that instruments chosen or developed frequently must be discarded subsequently because of pilot opinion, poor pilot performance, and/or system design changes.

The constraints put on spacecraft displays regarding weight, reliability, and reluctance to use untried techniques make the aerospacecraft display system a field full of new ideas, but the hardware being developed is mostly conventional (ref. 1). The main trend is to push conventional displays into as highly an integrated form as possible, without taking the major step of going to a single time-shared, general-purpose display and the consequent removal of the traditional maze of instruments. Single-purpose instruments and arrays of these instruments are, by their nature, limited in flexibility and provide no means of evaluating general-purpose display concepts. Thus it becomes imperative that a new capability be established to evaluate general-purpose display concepts utilizing the research simulator as a means of providing experimental data for a basis of comparison.

A POTENTIAL SOLUTION.—A solution to the above-mentioned display problems appears to lie in a new concept proposed in July 1962 by the author, who later found that a similar concept was being investigated at Wright-Patterson Air Force Base under contract to North American Aviation, Inc. (refs. 6 and 7).

This concept is based on the premise that it is possible to develop a programmable electronic display system which synthesizes desired instrumentation at the control-display interface. A pictorial diagram representing this concept in terms of simulator signal flow is shown in figure 1. It can be seen that such a system would operate in conjunction with a flight simulation computer and the simulated cockpit control-display interface. The synthesized displays would be driven dynamically in accordance with the flight equations as perturbed by pilot control inputs.

Assuming that such a display synthesis system could be developed, it would reduce a basically hardware problem to that of a software problem. If the design allowed for rapid and

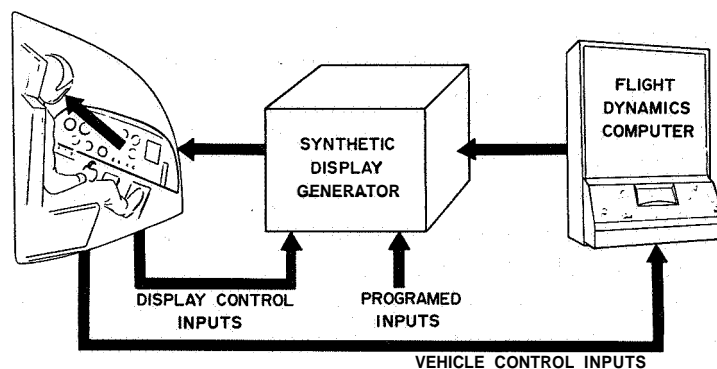


Figure 1.—Pictorial diagram showing simulator signal flow using programmed synthetic display generation.

efficient programing and was sufficiently universal in nature, it would have the potential for producing research simulator displays at lower cost, at higher speed, and at more advanced levels than conventional techniques will allow. For maximum efficacy such a display synthesizer should have many of the following characteristics:

- (1) Rapid and efficient programing with a minimal turn around time
- (2) Synthesis of desired instruments with the completed system requiring little or no new hardware design
- (3) Universal in nature; thus, capable of the synthesis of a wide spectrum of displays including electro-mechanical as well as electronic and electro-optical displays
- (4) Employing a combination of devices not exceeding the state-of-the-art and producing a feasible, reliable system
- (5) Capable of use at a central location with remotely driven displays, which are compatible with fixed-base and dynamic flight simulator cockpits
- (6) Compatible with flight simulation computers and associated trunking networks
- (7) Utilizing, if possible, new techniques being proposed for, and directly applicable to, next generation flight vehicles, thereby making it useful as a test bed as well as a simulation research tool

The remainder of this paper will be devoted to (1) a discussion of the technique chosen as a basis for synthetic flight display generation, (2) the description of a programmable display synthesis system utilizing this technique, and (3) the discussion of initial system performance.

SYNTHETIC DISPLAY TECHNIQUE CHOSEN

POTENTIAL TECHNIQUES.—Of the display techniques studied which are applicable to the synthetic generation of flight displays, those classes of displays known as programmed electronic displays afford the most promise for providing a repertoire ranging from simple electro-mechanical displays to sophisticated general-purpose displays (ref. 8). Most programmed electronic displays come under the category of computer-generated CRT displays which have the desired advantages of unrestricted display format, of good image quality, and of being programmable.

A major problem is encountered, however, in the area of programing, in attempts to apply existing computer-CRT displays to the task of flight display synthesis. The programing requirements for effective display can be very extensive, often running to many thousands of digital words (ref. 9). Specific disadvantages are encountered with the character and vector generation schemes generally in use with CRT displays (ref. 10). These disadvantages may be categorized as follows:

(1) Stylized display—Since all dynamic displays must be composed of alphanumerics and vectors, the visual image takes the form of rudimentary line drawings or stylized displays rather than continuous tone, photographic-type displays.

(2) Lengthy programs—Since all characters and lines in the dynamic portions and many times in the static portions of visual displays must be selected, positioned, and unblanked on an element-by-element basis, the computer word program can become quite lengthy.

(3) Regeneration rates—Since each element of the visual display must be manipulated individually and regenerated individually at a rate no less than 25 cps to avoid flicker, the digital word rate required for generation of complex displays can become high enough to prohibit interlacing the display program with other computer control and arithmetic sequences. In these cases a separate recirculating memory is required in the display console for regeneration of the display.

(4) Character change—Display symbols can be changed only through the substitution of new circuit modules, scanning tubes, or display tubes.

The disadvantages of these character generation schemes limit, but do not preclude, the direct use of existing computer-CRT combinations as a synthetic flight display generator. However, in an effort to implement a more easily programmed and a truly utilitarian flight display synthesizer a new means of dynamic pattern image generation was investigated. This dynamic pattern image generation method is described in the following two sections.

PRINCIPLE OF OPERATION.—The techniques selected for dynamic and static image generation and for overall display synthesis are based upon the principle illustrated by figure 2. This principle asserts that most desired flight displays are composed of static patterns and dynamic patterns which can be separated for photographic storage and, under the control of programmed instructions, machine dynamics inputs, and manned inputs, can be electronically recombined for composite, dynamic display. A programmed display synthesis technique based on this principle can best be described by the term, "electronic animation."

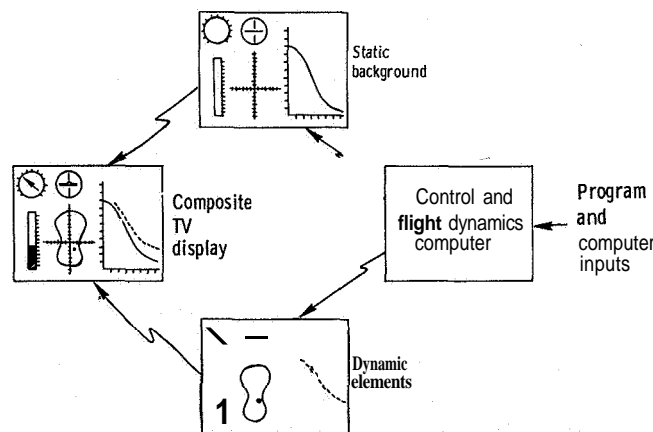


Figure 2.—Display synthesis principle—electronic animation.

RUDIMENTARY SYSTEM REQUIREMENTS.—The basic requirements for electronic animation consist of (1) the means for calling up static and dynamic display patterns from random-access film storage, (2) the means for electronically modifying scanned dynamic patterns (in accordance with flight dynamics equations) to convey motion, and (3) the means for combining static and dynamic portions of the display to form animated composites.

The basic system components required to achieve electronic animation are shown in the conceptual system diagram of figure 3. For the conceptual system, high resolution closed-circuit monochrome TV is chosen as the display means. This choice provides for flexibility and economy of display since closed-circuit TV displays are available in many configurations and are relatively inexpensive. In addition, the use of closed-circuit TV as the display means allows for use of the synthetic display generator at a central location with remotely driven displays, which can be made compatible with fixed-base and dynamic flight simulator cockpits.

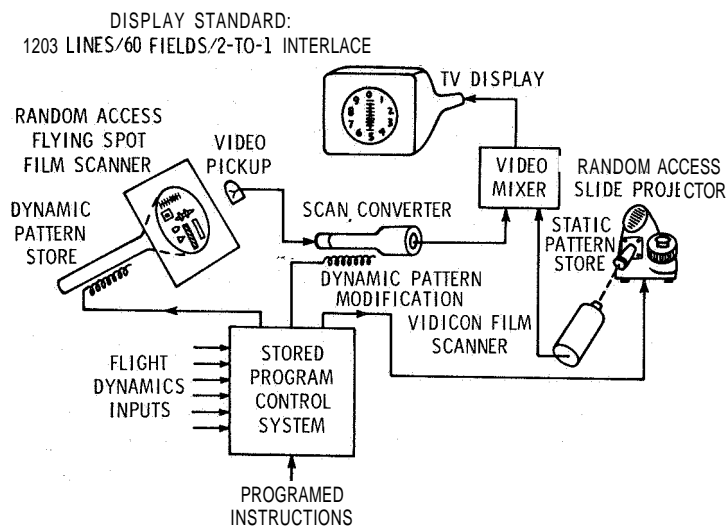


Figure 3.—Conceptual system diagram showing basic components required for electronic animation.

The use of the television display technique allows the combining of static and dynamic portions of the composite, animated display through simple video mixing. The video mixing, however, requires that both static and dynamic video picture information be in the closed-circuit television format of the display monitor, which is 1203 lines/60 fields/2-to-1 interlace. Thus, static pattern storage can be accomplished utilizing a random-access slide projector. Video pickup of static picture information is accomplished by a closed-circuit TV vidicon film scanner into which the static image is projected.

The dynamic (animated) portion of the composite display is got into a closed-circuit TV format through the use of the scan converter. The scan converter is necessary for this function since multiple dynamic patterns must be scanned and manipulated individually at the input to this interface and yet must be arrayed in the correct geometrical pattern and scanned as a whole at the output of the interface.

The choice of the dynamic image-generation method is by far the most important factor in the shaping of an electronically animated display system. This is because it is this method that sets the requirements for control sequences, storage capacity, regeneration rates, and programing. The dynamic image-generation method shown in rudimentary pictorial form in figure 3 is that of a random-access flying-spot film scanner.

The random-access flying-spot film scanner performs the function of calling up from photographic storage those patterns which are to be the dynamic portion of the composite, synthetic display. Since most flight display panels contain multiple dynamic display patterns, each of which must relate motion individually, the use of multiple pattern photographic storage and of individual pattern scanning with diminutive rasters is indicated (ref. 11).

The use of a raster-scan pickup for dynamic image generation implies the use of a raster-scan input to the scan converter interface, where the pickup and display (interface) raster are scanning in synchronism. Each pickup raster must be positioned to the appropriate dynamic element on the film store, and its corresponding display raster must be positioned to the appropriate display location in the scan converter at the appropriate time in the display composing sequence. Additionally, some means must be provided ~~for~~ animating ~~or~~ providing appropriate motion to each dynamic element of the display. This animation is provided in the conceptual diagram of figure 3 by electronic modification of the position, size, shape, and/or orientation of each display raster before it is written into transient storage in the scan converter. The electronic animation is accomplished in accordance with the appropriate sampled flight dynamics channel from the flight simulation computer.

The final requirement for implementation of a conceptual electronic animation system is for some form of program storage and control-sequence generation. This function can be performed by a small digital computer ~~or~~ by a special-purpose programable control unit as shown in figure 3.

DESCRIPTION OF EXPERIMENTAL HARDWARE

The synthetic flight-display generation equipment actually implemented at Langley Research Center is not described in detail, but rather it is described generally from the viewpoint of a potential user. Programing techniques, operational modes, and flexibility of the "electronic animation" technique are emphasized. Those readers requiring a more detailed description of the display hardware are referred to references 8 and 12.

GENERAL CONFIGURATION.—A photograph of the hardware developed to implement electronic animation techniques ~~for~~ research simulator display applications is shown in figure 4. Emphasis in the design of the system was on (1) reliability, through the use of solid-state components wherever possible; (2) maintainability, through the use of laboratory-size as opposed to microminiature construction; (3) programability, through the minimization of software and coded digital instructions; (4) universality, through the employment of rapidly changeable static and dynamic film-store patterns, the latter of which are electronically animated; and (5) flexibility, through the use of closed-circuit TV as the readout means.

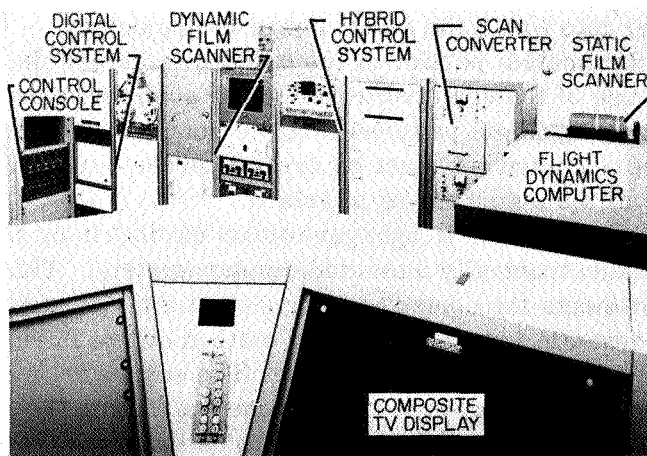


Figure 4.—Experimental synthetic flight-display generation equipment showing all major components.

The system design for the synthetic display generator is based on the conceptual system diagram of figure 3. The major subsystems involved are identified in the system photograph of figure 4.

The random access flying spot film scanner shown in figure 4 utilizes digitally controlled beam positioning and raster generation for access to film-stored dynamic elements for the synthetic display. The emphasis in the design of the scanner is on (1) providing high-speed random access to any pattern in the addressable transparency store matrix, (2) the scanning of the pattern in such a manner as to hold horizontal and vertical resolution constant, and (3) maximizing the beam utilization efficiency regardless of the size and shape of the scanned pattern. The scanner is designed for the scanning of 4- by 5-inch cut film, lantern slides, 2 1/4- by 3 1/4-inch cut film, and double-frame, 35-mm slides in conjunction with a precision detent mechanism for accurate subject positioning. Optical magnification and demagnification of scanned subject material is available over a total range of 5-to-1.

The video generated by the diminutive, random-access raster-scan format of the flying-spot scanner is not usable for direct display on a closed-circuit TV monitor. It must be scan converted to the specified closed-circuit TV standard of 1203 lines/60 fields/2-to-1 interlace for display of instrument dynamics. The scan-conversion system utilized is shown also in figure 4.

The scan-conversion system is capable of simultaneous recording and readout of dynamic pattern information through the use of two single-gun recording storage tubes. Read and write modes take place in different tubes alternately and independently, but in phase. The scan-conversion system's modes and beam positioning are digitally controlled by the programable control unit. Display dynamics are attained in the scan converted output by the dynamic modification and positioning of the diminutive input rasters, which are updated at up to 15 cps to prevent motion breakup.

A portion of the static-pattern film scanner is shown in the upper right hand corner of figure 4. Shown is the vidicon film camera into which 35-mm random access slide projectors project through an optical multiplexer. The random-access slide projectors are under the control of the programable control unit and can be controlled manually by either the test subject or an experiment controller. The drum slide holder has an average access time of 2 1/2 seconds between slides. The projectors can be sequenced digitally, however, with optical multiplexer shutter control to eliminate time lapses between slide changes. The two projectors are capable of holding 96 static patterns in random-access storage.

As shown in figure 4, the stored program-control system is broken up into a digital portion and a hybrid portion. The design of the digital portion of this subsystem is based on the utilization of program storage in a random-access magnetic-core memory. The memory is sequentially scanned in a recirculating manner to produce the control sequence of digital words necessary for display generation. The manner in which the parallel memory-word output is utilized for display-system control is based on the one-address instruction type of coding format used in digital computers in which a portion of the digital word contains the operation code and another portion of the digital word contains the operand address. This coding format was selected on the grounds of machine simplicity and simplicity of coding (ref. 13).

In addition to the memory, processing logic, and coded-instruction tape readers, the digital portion of the stored program control system contains decoding logic for the selection of static patterns stored in the random-access slide projectors. The hybrid portion of the stored program control system shown in figure 4 contains (1) a digitally controlled raster generator for the scanning of dynamic patterns, (2) four channels of digital-to-analog conversion for scanner beam positioning to programmed dynamic patterns and for scan-converter beam positioning to programmed display locations, (3) an analog-signal multiplexer for

sampling animation control signals from a flight dynamics analog computer, and (4) a raster modification and positioning unit for electronic animation of scanned dynamic patterns as they are being written into the scan converter.

A portion of the display console is shown in the foreground of figure 4. This console houses remote controls for the display system, the video mixer, and closed-circuit television monitors for the monitoring of static, dynamic, and composite display formats. Space is provided in the console for a future on-line compiler to facilitate programming which, while the system is being completed, is done directly in machine language.

Once the synthetic display generator is completed, a variety of closed-circuit television display techniques may be used at remote simulator cockpit sites. Figure 5 shows several of the television display techniques which could be made available. At present the display system is not tied in with a fixed-base or dynamic simulator; therefore, the only display means of figure 5 being employed is that of the television kinescope display. Figure 6 illustrates the two basic techniques which could be employed to channel closed-circuit television pictures to remote sites. Coaxial cable links are feasible at ranges of up to several miles. Microwave relay links could be used for longer ranges and for airborne simulations of advanced control and display systems. Work at Langley Research Center (ref. 14) has shown that such airborne simulations are feasible using standard IRIG telemetry channels for the transmission of sensor and pilot-control data for the simulated display system.

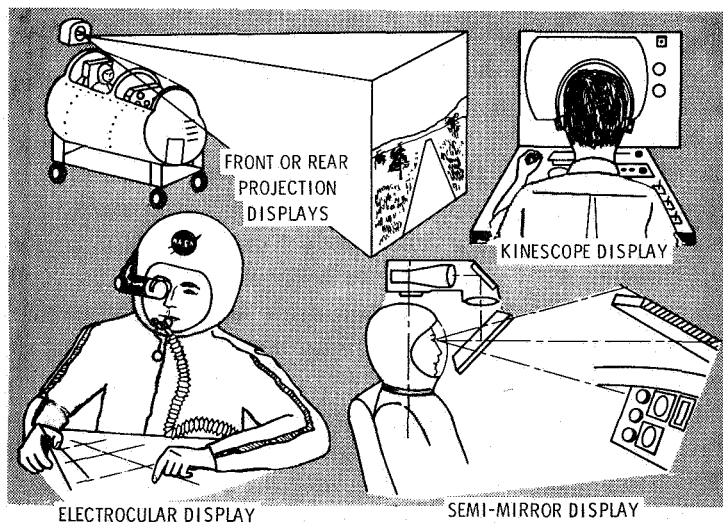


Figure 5.—Several closed-circuit television display techniques.

PROGRAMING AND SYSTEM OPERATION.—The goal in programing the electronically animated synthetic display generator is to allow the flight-control display-systems designer to proceed directly from his concepts in the form of cardboard mockups to simulated flight displays through the preparation of static pattern transparencies, a dynamic pattern transparency, programed punched tapes, and patch-board programs. The programing is thus composed of three phases: (1) technical illustration, (2) slide preparation, and (3) coded instruction and patchboard programing.

Technical illustration begins with a sketch of the desired instrument face or of the desired display formats (in the instance of a general-purpose, time-shared display). These sketches can be magnified or reduced and projected onto a drawing board by the equipment

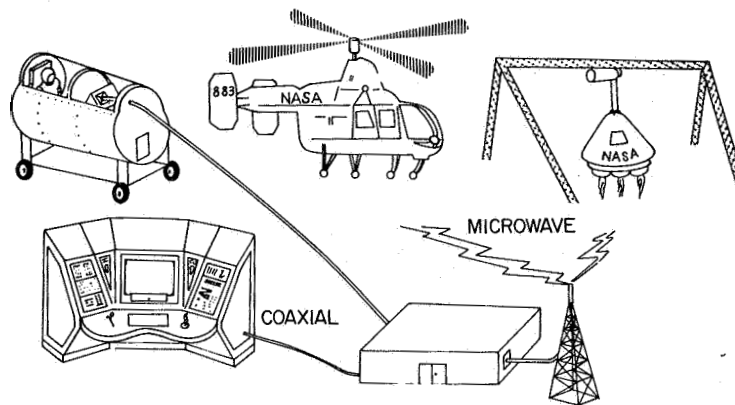


Figure 6.—Pictorial illustration of television distribution techniques for providing synthetic flight displays to remote dynamic or fixed-base simulators.

shown in figure 7(a) for fabrication of cardboard mockups. To increase the speed of the operation short-cut drafting aids such as tapes, shading sheets, and stickon symbols are used wherever possible. Mockups are separated into a dynamic pattern mockup and multiple static pattern mockups as is shown in figure 7(b).

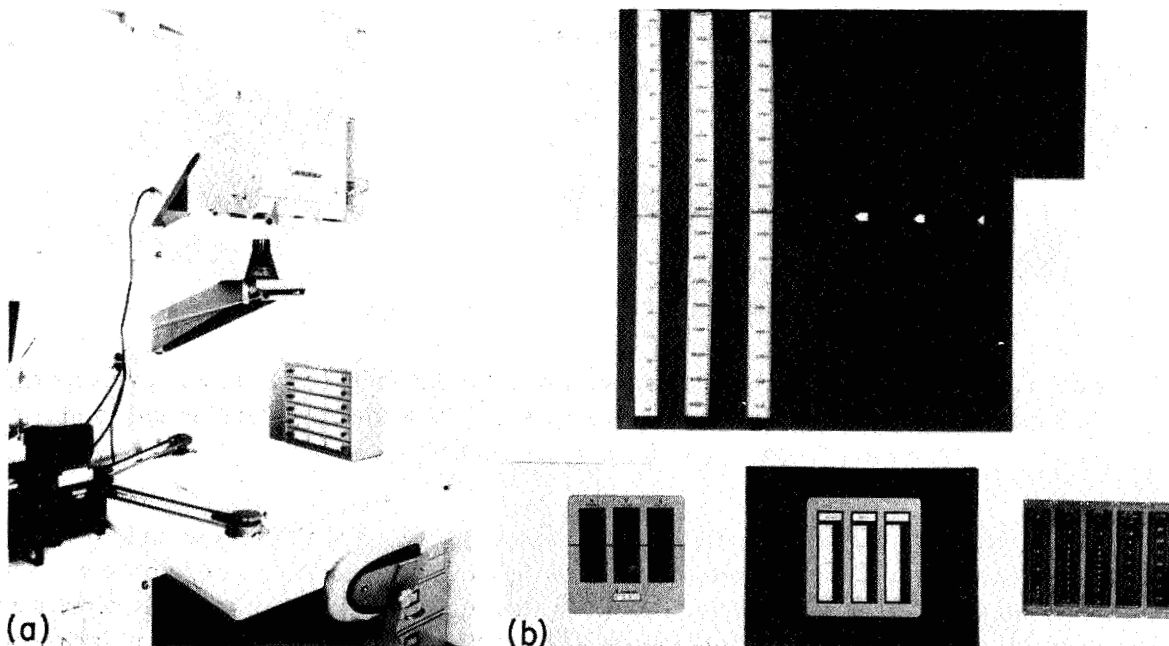


Figure 7.—Instrument mockup preparation. (a) Technical illustration aid; (b) typical dynamic and static patterns fabricated.

Slide preparation begins with the photographing of the static pattern and dynamic pattern cardboard mockups using the Polaroid industrial view camera shown in figure 8(a). High

quality transparencies can be obtained within 5 minutes using Land projection film types 46L or 146L. The resulting dynamic pattern transparency (3 1/4 by 4 in.) is used directly in one of the flying-spot-scanner film holders shown in figure 8(b). The static-pattern transparencies produced must be punched for a precision-registration slide mount for 35-mm film scanning. The precision-registration and punch equipment is shown in figure 8(c), and the precision 35-mm television slide mount is shown in figure 8(d).

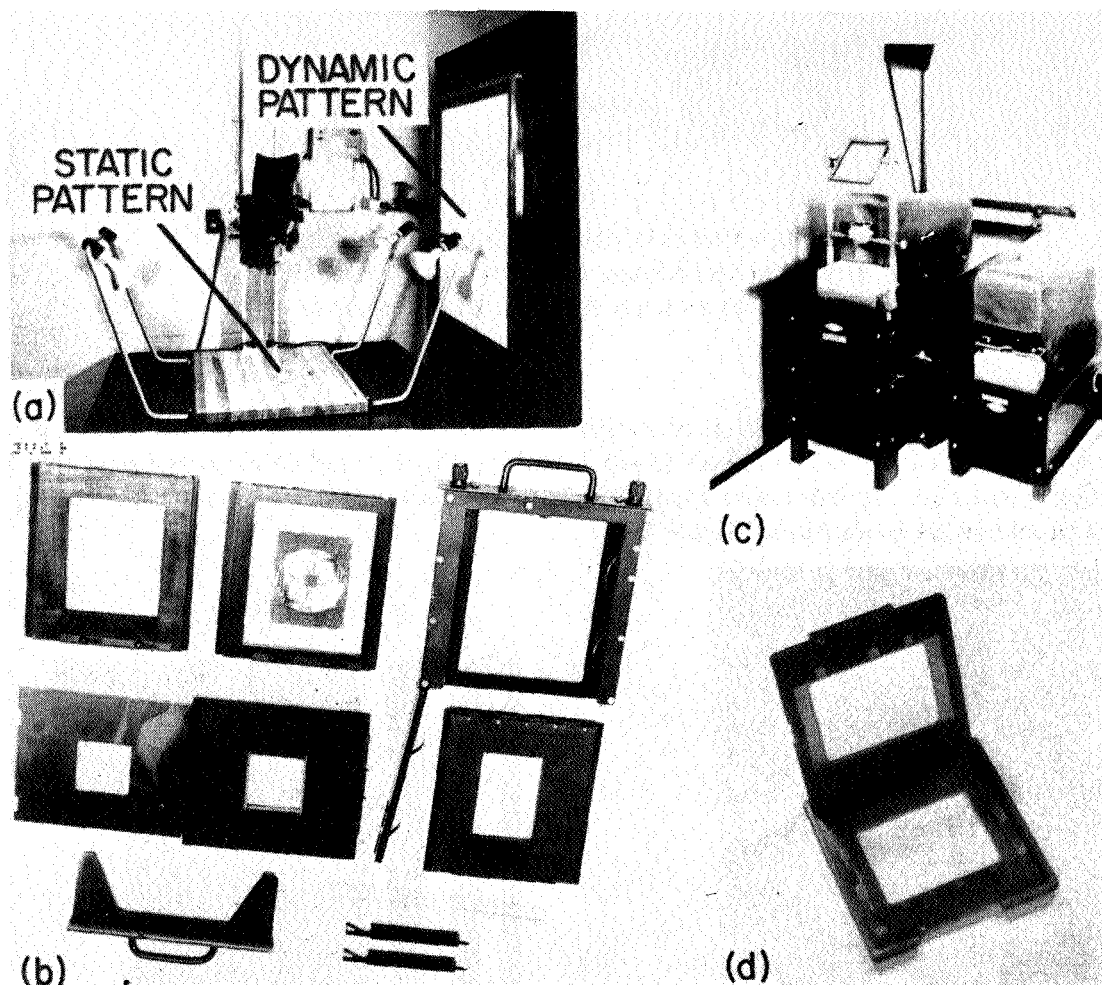
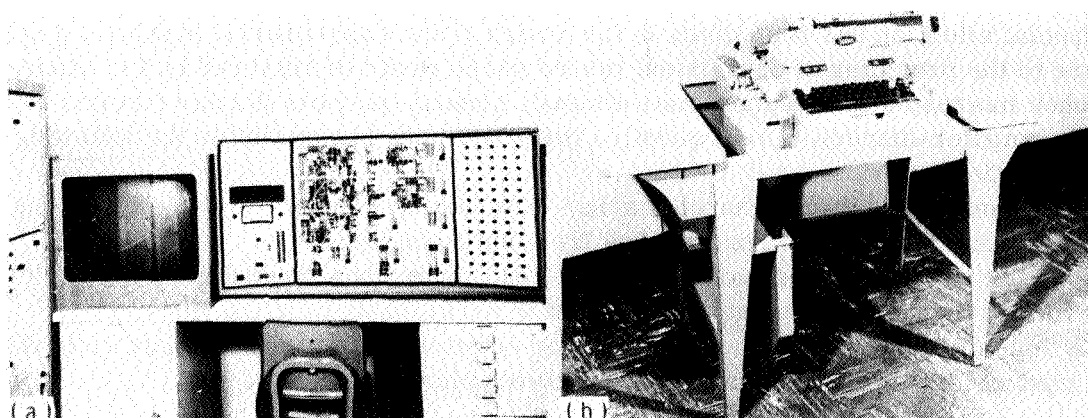


Figure 8.—Transparency-preparation equipment. (a) Industrial view copy camera; (b) dynamic pattern film holders; (c) precision TV slide registration unit and punch; (d) precision 35-mm television slide mount.

The final step in programming, the preparation of patchboard programs and coded instructions, is effected by the equipment shown in figures 9(a) and 9(b), respectively. The removable patch-panel analog computer shown in figure 9(a) is programmed for a simulation of flight dynamics. This 58-amplifier computer is suitable for small-scale simulations; however, for very complex problems a tie-in with a larger computer is necessary.

The tape punch shown in figure 9(b) prepares four-level tapes using the keyboard and will be capable of preparing eight-level tapes with external electrical inputs from an on-line compiler. Thus, the instructions for the stored program control system are loaded into its memory from the punched tape in the form of 24-bit words composed of six 4-bit bytes or three 8-bit bytes.



SUBSYSTEM BEING ADDRESSED	BIT INFORMATION NUMBERS																			CONTROL			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
STATIC PATTERN SELECTION	PROJECTOR A CODE									PROJECTOR B CODE									3 1 0 0				
	SIZE AND SHAPE CODE									WRITE OR ERASE CODE													
DYNAMIC RASTER SELECTION																			1 0 0 0				
DYNAMIC PATTERN SELECTION	X ADDRESS CODE									Y ADDRESS CODE									0 0 0 1				
	X ADDRESS CODE									Y ADDRESS CODE													
DISANIMATION CONTROL																			0 0 1 0				
ANIMATION CONTROL SELECTION	CONTROL CHANNEL CODE									ANIMATION FUNCTION CODE									1 0 0 1				

Figure 9.—Programed-instruction preparation equipment. (a) Flight dynamics analog simulation computer; (b) keyboard tape punch; (c) coding format.

The five basic types of control system information words utilized and the functions of each are illustrated in figure 9(c). Programming of the synthetic display system is greatly simplified through the use of only these five basic types of information words. The first type of information word, as can be seen from figure 9(c), is devoted to selection of static patterns from each of the two random-access slide projectors. Since the static portion of any given display format is unchanging, this instruction word is programed only once for any given display. The remaining four instructions must be programed once for each dynamic element which is to be animated. The programmer selects, with the first of these remaining four instructions, the size and shape of the raster for scanning a dynamic element or for erasure of dynamic elements stored on the scan conversion interface. He selects, with the second and third information words, the address of the dynamic element in film store and the display location for this dynamic element on the scan conversion interface, respectively. The final instruction listed in figure 9(c) is the selection of an animation control channel and an animation function (or functions) for that channel.

Animation functions which are to be available in the completed system are the following: (1) horizontal and vertical translations of dynamic pattern pickup raster, (2) horizontal and vertical translation of dynamic pattern display raster, (3) separate or simultaneous magnification of horizontal and vertical size of dynamic pattern pickup or display raster, (4) rotation of dynamic pattern pickup or display raster, and (5) oscilloscope-mode character or symbol writing directly onto the scan conversion interface.

Special attention has been given in the design of the experimental synthetic display hardware to the providing of operational modes which would be required by the simulation researcher interested in exploring time-shared, general-purpose display formats (which could be computer-driven). Consequently, the programmer has available the following time-shared operational modes:

- (1) Preprogramed time-shared display—The display format may be changed automatically as the simulated mission progresses from phase to phase.
- (2) Pilot adaptive time-shared display—Test subject has manual control over dynamic display format through a programmable pushbutton array.
- (3) Machine adaptive time-shared display—Control system can automatically display control parameters exceeding thresholds or having alarming trends.

PERFORMANCE ACHIEVED TO DATE

It is felt that a graphic demonstration of system performance is shown by photographs and motion picture films of typical synthesized displays. At the time of preparation of this paper the only animation functions available for synthetic display generation were translational and oscilloscope mode dynamics. Additional animation functions are being incorporated and a motion-picture film is being prepared in which the animation functions incorporated to date will be utilized in the generation of typical synthetic displays.

Figures 10(a), 10(b), and 10(c) show three types of vertical indicators synthetically generated by the use of translational electronic animation dynamics. These displays were photographed directly from a display system monitor and show the static and dynamic portions of the composite display and the composite display produced by the video mixing of the static and dynamic video signals. These displays were synthesized from the dynamic-pattern mockup and the three static-pattern mockups shown in figure 7(b). The displays shown are fully animated with no ambiguities or discontinuities. For instance, in the three-window vertical tape display the static pattern is seen to be a window with a lubber line. The dynamic pattern is a vertically oriented tape which appears to move up or down within the confines of the window according to the dynamics of a flight-control parameter. Short-term stability (over a 1-hour period) for the dynamic elements of typical displays, such as those shown in figures 10(a), 10(b), and 10(c), is such that changes in position are less than 0.5 percent of full-scale value.

System performance in the area of programing complexity is indicated by the program length, or number of digital words, required for the generation of specific displays. For those displays shown in figures 10(a), 10(b), 10(c), the program lengths are 36, 18, and 18 digital words, respectively. These program lengths are shorter than those required with conventional computer-CRT stylized displays, or specifically, more than a factor of ten shorter in the instance of the vertical tape display (ref. 12). It is felt that the shorter program length of many displays will ease the burden on the programmer.

The electronic animation technique at this stage of its development is not without its limitations. Resolution in the static portions of displays is quite good; however, resolution in the dynamic portions of displays requires some improvement. Some distortion is produced in dynamic portions of the display due to digital noise and a skewing of dynamic pickup rasters. Some flicker is evident in dynamic portions of the synthesized display due to the use of a two-tube scan conversion system. These limitations are in the process of being minimized. In addition, the capabilities for the synthesis of digital displays and displays having rotational dynamics are being incorporated.

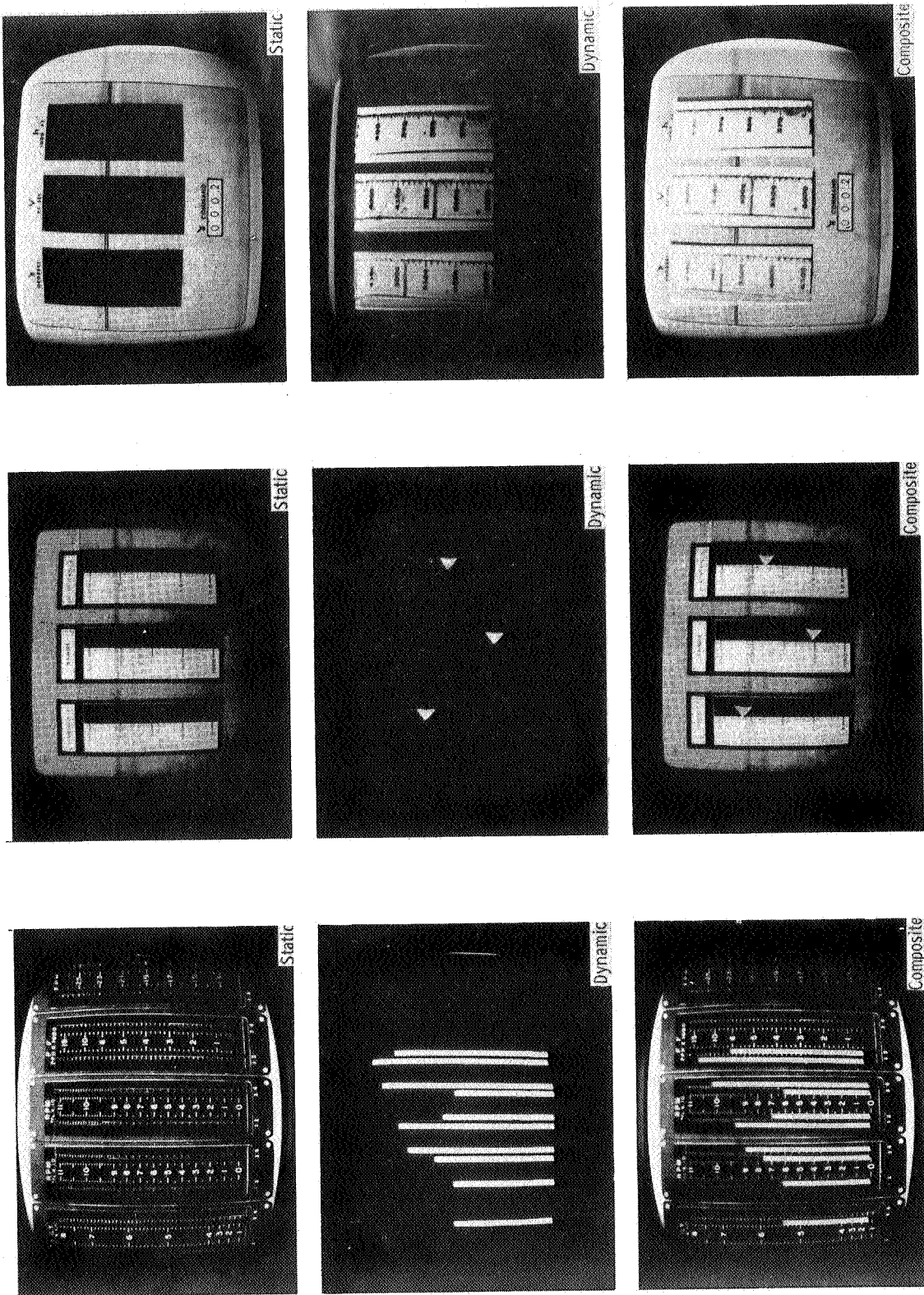


Figure 10.—Three types of vertical indicator displays photographed directly from a 21-inch monitor showing static, dynamic, and composite pictures. (a) Dynamic vertical bar indicator display; (b) dynamic vertical pointer indicator display; (c) dynamic vertical tape indicator display.

CONCLUSIONS

The research and development required to implement the system described has established the feasibility of rapidly programing a portion of the displays generally encountered in simulation and presenting them at the display interface via closed-circuit television.

In addition, this work has developed a man-machine communications research tool providing for a systems engineering approach to flight display panel design through the programmed synthesis of electronically animated flight displays. This research tool provides for the study of time-shared, general-purpose display concepts through its basic time-shared control modes.

A new concept has been investigated and proven feasible—that of electronic animation. The advantages which this concept can have in terms of programing complexity, display format changeability, and image quality for synthetic flight display generation in simulators have been pointed out.

In addition, electronic animation may have potential as an avionics display technique, either for simulation utilizing a ground-to-air microwave link or for self-contained general-purpose, computer-generated display. The stored program generation technique lends itself nicely to interlacing the display program with the control and arithmetic sequences of a central avionics computer since so few digital words are required. Electronically animated television displays are seen as a special advantage in a space vehicle such as a large space station, which would require integrated displays at multiple locations throughout the vehicle. However, much work remains to be done, both from a hardware and a human factor's standpoint, before a specific system design could be recommended.

REFERENCES

1. LaFond, Charles D.; and Pay, Rex: Special Report: Advanced Displays. Missiles and Rockets, Oct. 1964, pp. 27-47.
2. Anon.: DAEMON, An Advanced System for Data Management in Manned Space Vehicles, Document No. 2080B, American Bosh Arma Corp.
3. Shulman, A.: Display and Control in Manned Space Vehicles. Paper presented at 8th Annual Meeting of the American Astronautical Society (Washington, D. C.), Jan. 1962.
4. Koppa, R. J.: A Survey of Current Display and Control Concepts. Rep. no. 00.308, LTV Astronautics Div., Mar. 1964.
5. Buyan, J. R.: Manned Military Spacecraft Displays. Aerospace Corp. (Contract No. AF04(695) - 269), Jan. 1964.
6. Anon.: Cathode-Ray Tube Instrument Synthesis System. Tech. doc. rep. no. AMRL-TDR-63-84, Wright-Patterson Air Force Base, Oct., 1963.
7. Anon.: A Proposal for a Cathode Ray Tube Instrument Simulator. No. DNO 63-647, North American Aviation, Inc., Oct. 1963.
8. Hatfield, Jack J.: The Investigation and Development of a Programable Display Synthesizing System for Man-Machine Communications Research. Master's Thesis, Univ. of Va., Aug. 1966, pp. 33-56.
9. Corbin, Harold S.: A Survey of CRT Display Consoles. Control Engineering, Dec. 1965, pp. 77-81.
10. Boyd, Sherman H.: Digital-to-Visible Character Generators. Electro-Technology, Jan. 1965, pp. 77-88.
11. Jones, E. D.: Character Generator for Digital Computers. Electronics, **Feb.** 12, 1960, pp. 117-120.

12. Hatfield, Jack J.: A Programable Display Synthesizing System for Man-Machine Communications Research. Technical Session Proceedings , 7th National Symposium on Information Display (Boston, Mass.) , Oct. 18-20, 1966.
13. Grabbe, E. M., et al.: Handbook of Automation, Computation, and Control. Vol. 2, Computers and Data Processing, John Wiley & Sons, Inc. , 1959, p. 2-08.
14. Gracey, William; Sommer, Robert W.; and Tibbs, Don. F.: Evaluation of a Cross Pointer-Type Instrument Display in Landing Approaches With a Helicopter. NASA TN D-3677, 1966.

PRECEDING PAGE BLANK NOT FILMED.

II. DESCRIBING FUNCTION MODELS

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100.

101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200.

7. Inflight and Ground Simulation Measurements of Pilot Transfer Characteristics in the Compensatory Roll Tracking Task*

F. D. Newell

Cornell Aeronautical Laboratory, Inc.

This paper presents a review of a measurement experiment in which pilot transfer characteristics have been determined for a random-appearing forcing function in both ground-based simulation and actual flight. Three experienced test pilots were the subjects. The task was compensatory tracking in roll to maintain zero bank angle in the presence of small bank angle disturbances, of which the largest were approximately $\pm 8^\circ$. Two simple single-degree-of-freedom controlled elements were "flown" in the ground-based simulator only, and two six-degree-of-freedom airplane-like configurations with excellent handling qualities were flown both in the simulator and in actual flight. Companion work, not reported herein, has been accomplished by the NASA Flight Research Center with a contact analog, ground-based simulator.

Approximately 20 years ago, Tustin began the field which endeavors to describe the human controller sufficiently well mathematically to permit analysis of combined human-vehicle dynamic systems. Since then, the field has become strongly concerned with the description and analysis of the human pilot of aircraft for primarily compensatory tasks.

To give a brief outline, in the 1950's Elkind extended the knowledge of the field, Campbell and Eakin successfully extracted human transfer characteristics from flight data, and Hall, Seckel, McRuer, and Weir extracted human transfer characteristics from comparative flight and ground simulator data. Since then, McRuer, Graham, Krendel, and Reisener have reported an exceptionally thorough study and analysis of ground simulator data and have given pilot models which range from the very simple to the very sophisticated for compensatory tracking tasks. A logical continuation of these efforts has been to apply the knowledge obtained from them to a well conceived experiment in which inflight measurements of the human transfer characteristics are made. Such a measurement experiment and the results obtained from it are the subject of this paper.

The design used for the experiment was developed to include flight and ground measurements of the human operator in the same cockpit for comparable ground and flight controlled elements; to obtain ground simulator measurements of the pilot in a loop with simple, single-degree-of-freedom controlled elements which have been studied in the past; and to obtain an assessment of intrapilot variability.

*This work was sponsored by the U. S. Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio, under Contract AF33(615)-3605.

For the work reported, the Air Force variable stability T-33 was used as both a ground-based simulator and an in-flight simulator. The experiment was done in cooperation with NASA Flight Research Center to include comparisons among T-33 ground and flight simulation and ground, contact-analog simulation. Only the T-33 portion of the experiment is documented herein.

The only task performed was a compensatory, roll tracking task with maximum bank angles of approximately $\pm 8^\circ$. All of the ground-based flying was performed "under the hood" by reference to flight instruments. A hood was not used in the actual flight phase, but one configuration was flown by reference to the flight instruments, and the other was flown by visual reference to the natural horizon. For the ground simulation the two simple controlled elements were single degree of freedom in roll: one being an integrator and the other being divergent and difficult to control. The remaining two ground simulation configurations and the flight configurations were six-degree-of-freedom configurations with excellent handling qualities.

This paper presents discussions of the pertinent conditions under which the data were obtained and the techniques of computation used for obtaining the transfer characteristics of the three pilots who participated. The paper discusses the experiment, the equipment and the data in that order.

SYMBOLS

| | |
|---------------|--|
| i | tracking task, sum of 10 sine waves input |
| i' | tracking task, modified sum of 10 sine waves input |
| K | general gain constant |
| K_c | controlled element gain constant |
| K_p | pilot gain constant |
| s | Laplace operator |
| Y_c | controlled element transfer function |
| Y_p | pilot random-input describing function |
| $Y_p Y_c$ | open-loop random-input describing function |
| δ_{AS} | aileron stick displacement |
| ζ_d | Dutch roll damping ratio |
| ζ_{SP} | short period damping ratio |
| τ | general time constant |
| τ_R | roll mode time constant |
| Φ_{ic} | system input—pilot output cross power spectrum |

| | |
|------------------|---|
| Φ_{ie} | system input—system error cross power spectrum |
| ϕ | airplane bank angle |
| $ \phi/\beta _d$ | ratio of roll amplitude to sideslip amplitude at the Dutch roll frequency |
| ω_d | Dutch roll frequency |
| ω_{SP} | short period frequency |

THE EXPERIMENT

There are several purposes which are relevant to the determination of the design of the experiment. They are, essentially in the order of importance, as follows:

- (1) To instrument the variable stability T-33 for obtaining data, in both ground-based simulator and actual flight environments, from which human-pilot transfer characteristics can be determined in the same vehicle, for the same pilots, and for the same task.
- (2) To measure human-pilot transfer characteristics in the variable stability T-33 as a ground-based simulator for controlled elements which have been investigated in the past by other experimenters for the purpose of comparing results with past experiments.
- (3) To compare pilot variability among and within themselves.
- (4) To investigate the stationarity assumption which is fundamental to the data analysis technique.

These purposes are incorporated in the design of the experiment as shown in figure 1, which displays the compensatory task investigated, and in table 1, in which the numbers in the blocks represent the number of usable data runs obtained for each pilot for each controlled element.

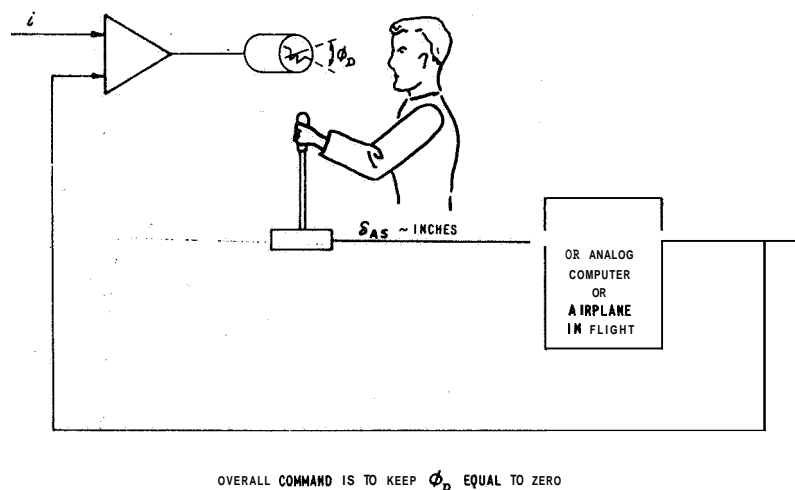


Figure 1.—Overall compensatory roll tracking loop.

The original intent was to obtain 10 runs for pilot A for each configuration, to obtain 10 runs for each pilot for configuration A-2, and to obtain 3 runs for each remaining configuration for pilots B and C. However, some data runs were lost as is indicated by the numbers in the table.

| Pilot | | | | | | |
|-------|-------------|---------------------------|-----|------|-------------|------|
| | T-33 ground | | | | T-33 flight | |
| | K/s | $\frac{K}{s[s-(1/\tau)]}$ | A-2 | A-2* | A-2 | A-2* |
| A | 10 | 10 | 3 | 10 | 8 | 10 |
| B | 3 | 3 | 10 | 3 | 10 | 3 |
| C | 3 | 3 | 8 | 3 | 10 | 3 |

For the K/s controlled element, the value of K in terms of roll rate per aileron stick deflection is 50 (deg/sec)/inch. The aileron stick spring rate is 2.5 lb/in.

For the $\frac{K}{s[s-(1/\tau)]}$ controlled element, the value of K in terms of aileron stick deflection is 120 (deg/sec)²/inch, the value of τ is 1 second, and the aileron stick spring rate is 2.5 lb/in.

No longitudinal control task was mechanized for either the K/s or the $\frac{K}{s(s-1/\tau)}$ controlled elements. The pilots' control task consisted of using the ailerons to minimize bank-angle excursions from zero. No rudder was required with either of these simple, single-degree-of-freedom controlled elements.

The A-2* and A-2 configurations were full six-degree-of-freedom mechanizations of the airplane equations of motion. The only difference between A-2* and A-2 was the manner in which the tracking task was introduced into the system. For A-2*, the tracking task was applied directly to the attitude indicator as indicated in figure 2. But for the A-2 case, the tracking task was introduced through the ailerons as depicted in figure 3. The complication which necessitates the signal, i' , as shown in figure 3, is that the tracking task which is displayed on the attitude indicator must correspond for both the A-2* and A-2 configurations. This required correspondence is accomplished by preshaping the input to the ailerons by the inverse of the bank angle per aileron deflection transfer function so that the actual bank angle which results from aileron motion is proportional to the tracking task i . The tracking task i , which is made up of 10 sinusoids, is called "the sum of 10 sine waves input" and the tracking task i' is called "the modified sum of 10 sine waves input."

So far as the pilot is concerned, the A-2* and A-2 configurations are the same on the ground-based simulator, but in actual flight, there is a difference between them. This difference is with respect to the bank angle which is displayed on the attitude indicator and the actual bank angle of the airplane. In flight with the A-2* configuration, the actual bank angle of the airplane is solely a result of the pilot's aileron stick inputs whereas the displayed bank angle is the sum of the actual bank angle and the tracking task. However, with the A-2 configuration in flight, the motion of the airplane is a result of both the pilot's aileron stick inputs and the tracking task input, and the bank angle displayed on the attitude indicator is identically the bank angle of the airplane. Thus the A-2 configuration can be flown in flight by visual reference to the natural horizon, and the pilot will see the tracking task input. However, if the pilot were to fly the A-2* configuration in real flight by visual reference to the natural horizon, then he would not see the tracking task. An additional difference between the two configurations in flight is that the nonvisual motion cues for A-2 are in harmony with the motion of the airplane, whereas for the A-2* Configuration, these cues are not necessarily in harmony with the motions displayed on the attitude indicator.

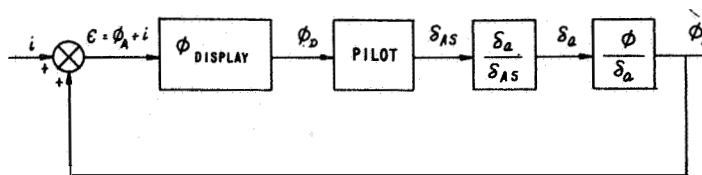


Figure 2.—Block diagram for use of unshaped input

$$\text{for } \frac{\phi_A}{\delta_{AS}} = \frac{K}{s}, \frac{K}{s\left(s - \frac{1}{\tau}\right)}, \text{ and } A-2^*.$$

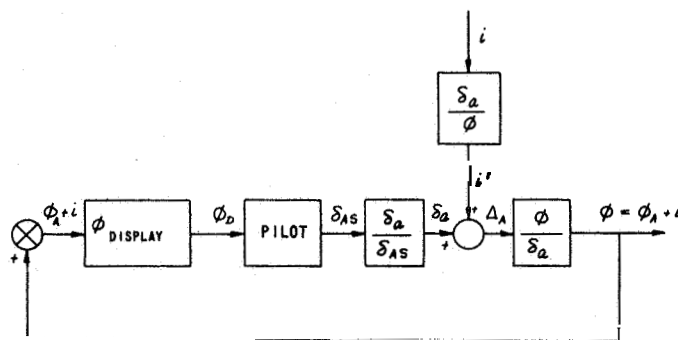


Figure 3.—Block diagram for use of shaped input

$$\text{for } \phi_A = A-2. \text{ The output is the product } \Delta_A \left(\frac{\phi}{\delta_a} \right) = \frac{\delta_{AS}}{\left(\delta_a + i \frac{\delta_a}{\phi} \right) \delta_a} \phi = \phi_A + i, \text{ and the figure may be re-drawn to appear as figure 2.}$$

The pilot was aided with the A-2 configuration in flight by four grease-pencil lines on the windscreen which helped to define small roll angle displacements and rates from steady wings-level flight. The placement of these lines is shown in figure 4.

For both the A-2 and A-2* configurations, the lateral-directional characteristics are the same and the aileron stick spring rate is 2.5 lb/inch. The overall transfer function is of the form

$$\frac{\phi}{\delta_{AS}} = \frac{K \left(s^2 + 2\zeta_{\phi} \omega_{\phi} s + \omega_{\phi}^2 \right)}{s \left(s + \frac{1}{\tau_R} \right) \left(s^2 + 2\zeta_d \omega_d s + \omega_d^2 \right)}$$

Also, the longitudinal characteristics are the same for both configurations and they are, as well, the same for flight as for ground-based simulation. These characteristics are presented in table 2. The aileron characteristics are such that aileron motion causes only a very small sideslip disturbance, and consequently the pilots did not use the rudder. Furthermore, the numerator of the bank angle per aileron deflection transfer function was adjusted to cancel the Dutch roll quadratic term in the denominator of this transfer function. Thus, the bank angle per aileron deflection transfer functions for the A-2* and A-2 configurations are shown in table 2 as second-order functions.

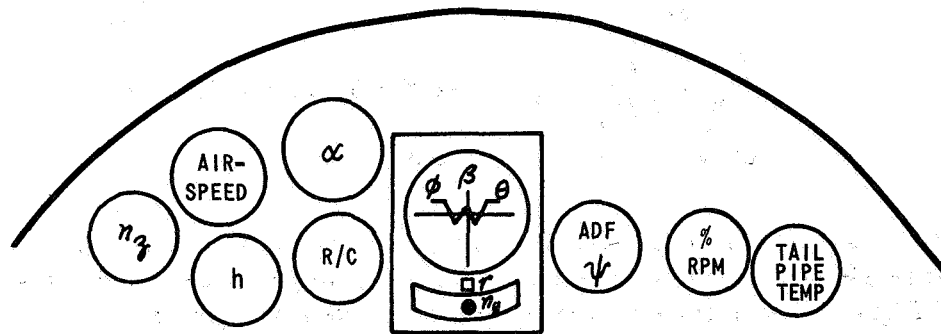
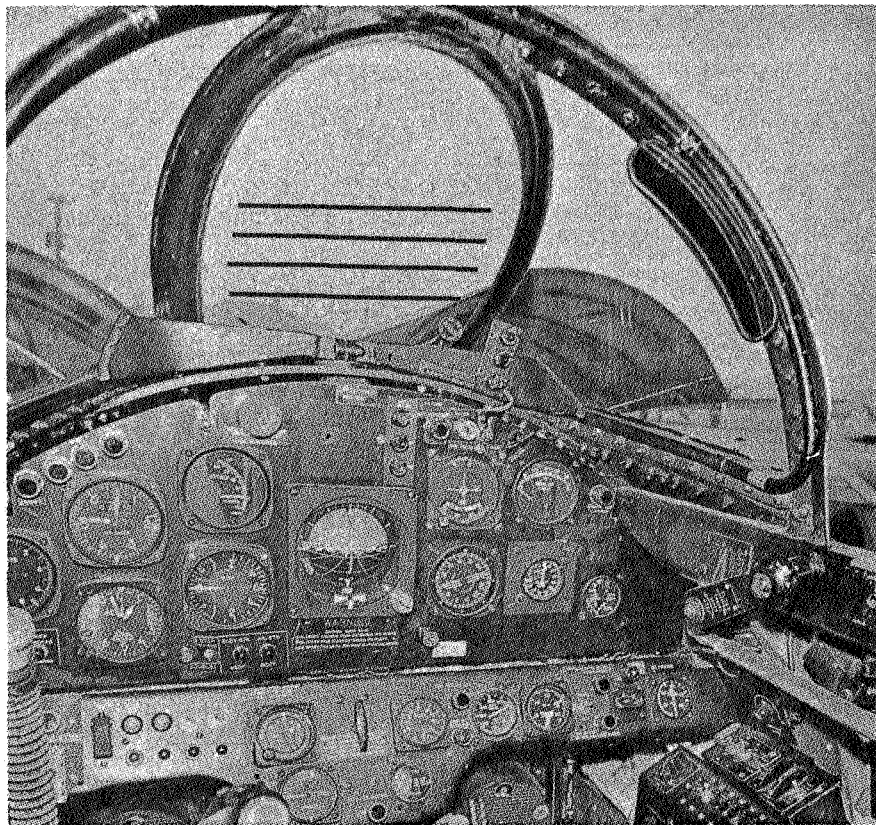


Figure 4.—Inflight simulator cockpit.

The differences in K_c , as shown in the table, for ground and flight were unintentional. However, the roll mode was so simple to fly that the only difference that is expected in the results is a different pilot gain, K_p . The pilots did not mention that they noticed a difference in K_c between ground and flight. There also was a difference in the phase of $|\phi|\beta|_d$ between ground and flight that is of little consequence in this experiment because the pilots' use of ailerons excited very little sideslip. The maximum excursions of β noted on several runs were a $\sim 3^\circ$.

The tracking task for the A-2 and A-2* configurations, both in flight and on the ground, is primarily a compensatory bank angle task with very little longitudinal control necessary to maintain an altitude of 23 000 feet and an indicated airspeed of 250 knots. No purposeful perturbation of the pitch attitude was introduced into the experiment.

TABLE 2.—A-2 AND A-2'' CHARACTERISTICS

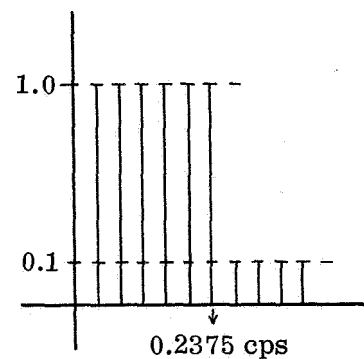
| Parameter | Ground | Flight |
|---|--|--|
| ϕ ----- | K_C
$s \left(s - \frac{1}{\tau_R} \right)$ | K_C
$s \left(s - \frac{1}{\tau_R} \right)$ |
| K_C , (deg/sec) ² /in. --- | 60 | 45 |
| τ_R , sec ----- | 0.35 | 0.35 |
| $ \phi \beta _d$ ----- | 0.6 | 0.6 |
| ω_d , rad/sec ----- | 2.28 | 2.30 |
| ζ_d ----- | 0.11 | 0.09 |
| Longitudinal | | |
| ω_{SP} , rad/sec ----- | 2.94 | 2.94 |
| ζ_{SP} ----- | 0.54 | 0.54 |

The tracking task signal time history has a Gaussian-like amplitude distribution and the signal was generated by summing 10 sine waves. The signal is in accordance with that shown in figure 20 of reference 1 and is defined to have a bandwidth ω_i of 1.5 rad/sec. The frequency content of the tracking task signals is included in table 3.

TABLE 3.—FREQUENCY CONTENT OF TRACKING TASK SIGNALS

| Frequency | |
|-----------|--------|
| rad/sec | cps |
| 0.157 | 0.0250 |
| 0.262 | 0.0417 |
| 0.393 | 0.0625 |
| 0.603 | 0.0958 |
| 0.969 | 0.1543 |
| 1.492 | 0.2375 |
| 2.539 | 0.4042 |
| 4.032 | 0.642 |
| 7.566 | 1.2047 |
| 13.797 | 2.196 |

RELATIVE AMPLITUDE FOR
TRACKING TASK, i



Although the sum of 10 sine wave signals appear to the pilots as random signals, they are actually periodic with a period of 4 minutes. However, each run by any pilot for each

configuration was of 4 minutes' duration, and there is no evidence of any sort that the pilots learned the signals in any way whatsoever.

No attempt was made to randomize the order in which the pilots flew the several configurations, either on the ground or in the air. This approach was demanded by the short time that the participating pilots were available. Whenever a configuration was set up, it was flown until all of the required runs for each pilot were obtained. There was no planned sequencing of the pilots, and each pilot did three to six 4-minute runs with a configuration on the ground-based simulator before another pilot flew. In flight, only three runs could be recorded per flight and, hence, each pilot did three 4-minute runs of a configuration per flight.

Each pilot was given time to practice with a configuration before data runs were made. However, none of the pilots did much practicing although they all did the most practicing with the $\frac{K}{s(s-1/\tau)}$ configuration because it is divergent and requires the pilot's complete attention. The little amount of practicing that was done means that the data presented herein represent actual pilots doing a representative piloting job.

THE EQUIPMENT

The equipment used in this experiment included a variable stability T-33, three PACE TR-10 and one TR-20 analog computers, and PEMCO FM tape, playback, and record units.

The instrument displays that were used in the T-33 are shown in figure 5 for the ground-based simulator. The instruments that were active for the ground-based simulation were: the normal acceleration η_z ; airspeed; altitude h ; angle of attack α ; rate of climb (R/C); the Lear all attitude indicator with bank angle ϕ , sideslip β , pitch angle θ , and yaw rate r ; the compass card of the ADF (RMI) indicator; and the percent rpm indicator. For the inflight simulation, all of the instruments on the panel as shown in figure 4 were operable, but the pilot concentrated on using the same instruments that were used in the ground-based simulation.

The analog computers used in conjunction with the T-33 for the ground simulation are PACE TR-10 and TR-20 transistorized computers produced by Electronic Associates, Inc. Three TR-10's and one TR-20 were used.

The cockpit control characteristics of the T-33 are displayed in figures 6, 7, and 8. The aileron stick characteristics are the same for all configurations. Although the elevator stick and rudder pedal feed systems were operative, there were no elevator and rudder inputs for the K/s and $\frac{K}{s[s-(1/\tau)]}$ configurations. It is noticed from figure 7 that the aileron stick has a hysteresis loop between stick force and stick position. The hysteresis was noticeable to all three pilots, but it was not considered to be objectionable.

Verification of the individual configurations was accomplished by analyzing transient responses to control impulse and step inputs, and by spectral analysis of the tracking records. The transfer functions obtained from this latter analysis are shown in figures 9, 10, and 11. A transfer function for the A-2 configuration computed from the recorded data is not shown since the total input ($\delta_{AS} + i'$) to Y_c was not recorded. However, the A-2 configuration transfer function is necessarily the same as that shown for the A-2* configuration. Precisely the same analog computer settings and variable stability system gains were used for the A-2 and A-2* configurations. It is seen from the transfer functions for the A-2* configuration that the Dutch roll mode is absent.

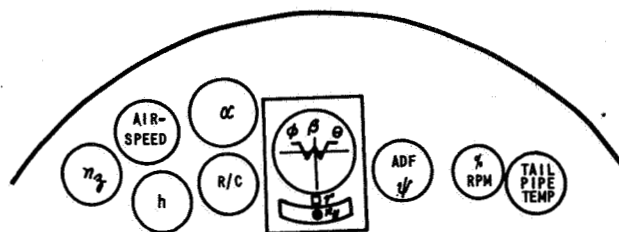
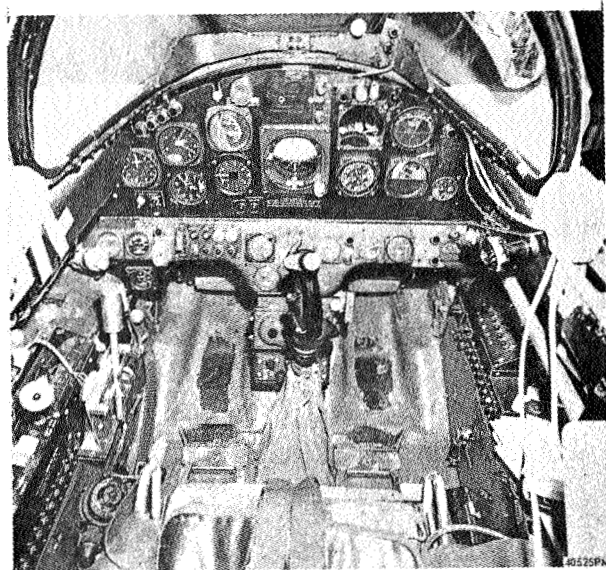


Figure 5.—Ground simulator cockpit.

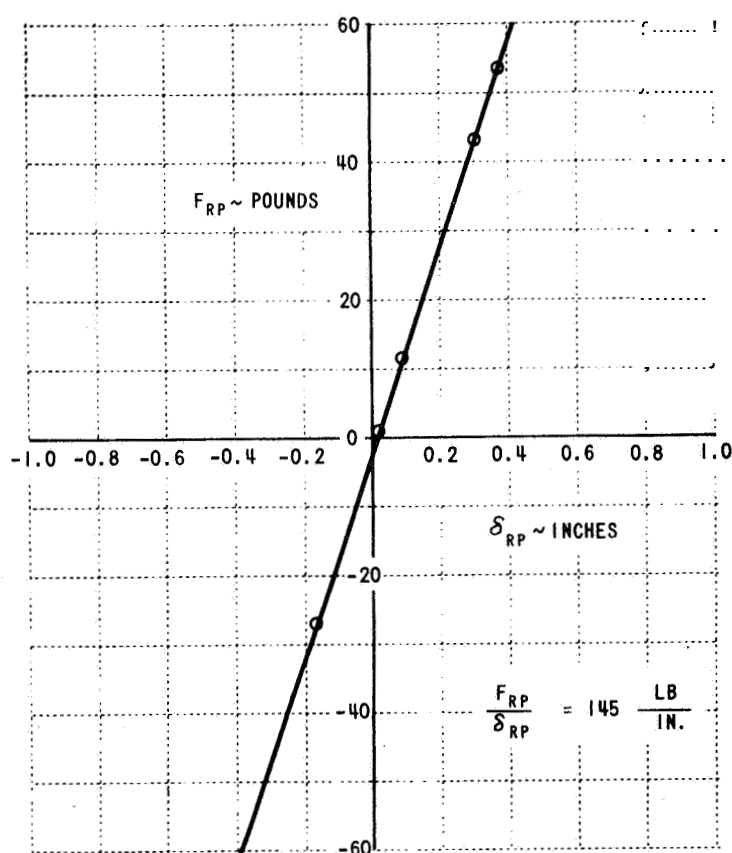


Figure 6.—Rudder pedal force versus displacement.

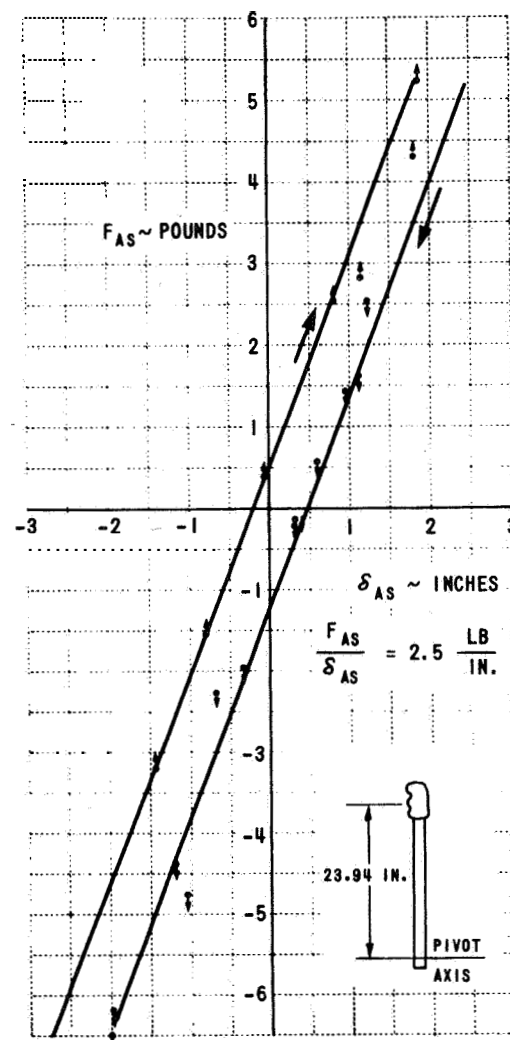


Figure 7.—Hysteresis loop of aileron stick force versus aileron displacement.

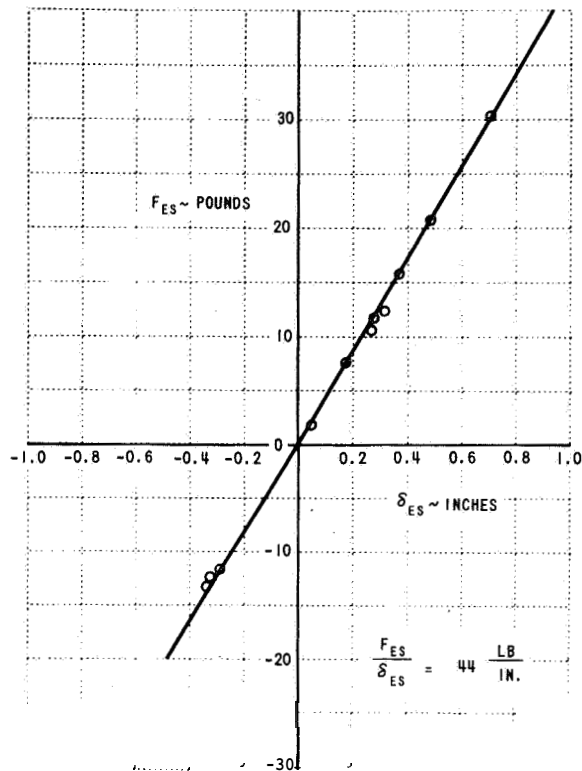


Figure 8.—Elevator stick force versus displacement.

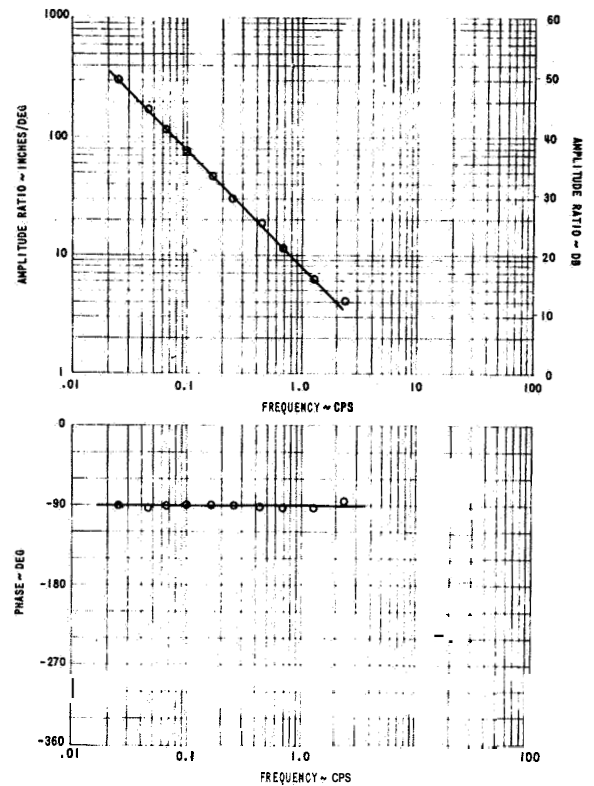


Figure 9.— Y_C for controlled element K/s .

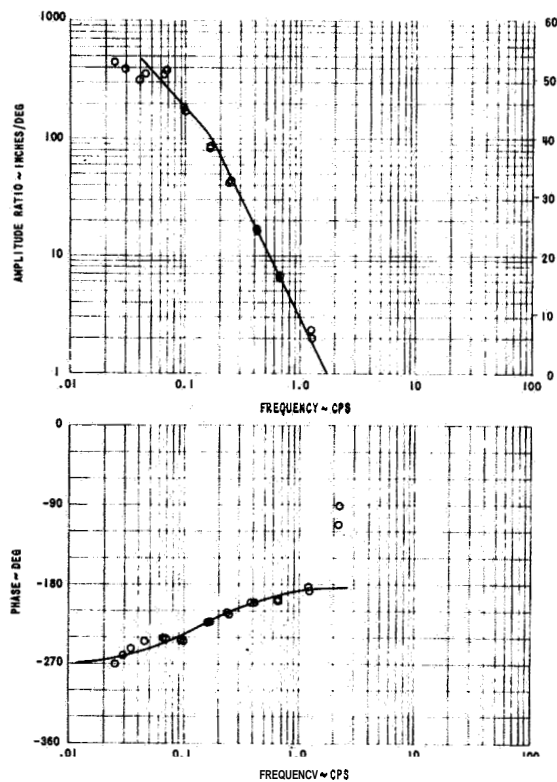


Figure 10.— Y_C for controlled element K/s ($s-1/\tau$).

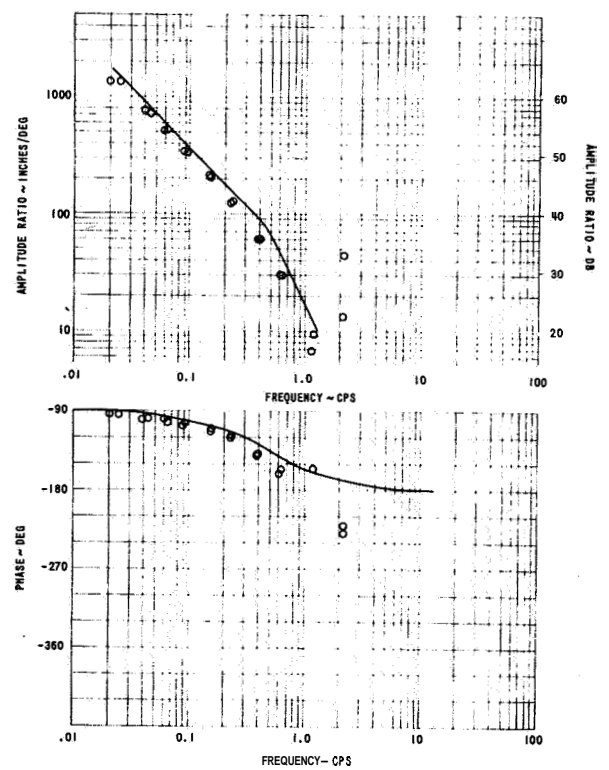


Figure 11.— Y_C for controlled element $A-2^*$ ground.

THE DATA AND DATA ANALYSIS

The data recorded on the FM tape were the tracking signal i or i' , the aileron stick displacement δ_{AS} , the airplane bank angle ϕ_A , and the bank angle displayed on the attitude indicator ϕ_D . These data were digitized in an analog-to-digital converter to fixed-point digital data. These digital data were analyzed by digital computers for the pilot random-input describing function Y_p , the controlled element transfer function Y_c , and the open-loop random-input describing function $Y_p Y_c$ according to power spectral density ratios.

That is ,

$$Y_p = \frac{\Phi_{ic}}{\Phi_{ie}} = \frac{\Phi_{i\delta_{AS}}}{\Phi_{i\phi_D}}; Y_c = \frac{\Phi_{i\phi_A}}{\Phi_{i\delta_{AS}}}; Y_p Y_c = \frac{\Phi_{i\phi_A}}{\Phi_{i\phi_D}}$$

where the symbol Φ_{ab} represents the cross spectral density between data trace "a" and data trace "b".

Because the digital program for computing the spectral densities had been developed for continuous random signals whereas the tracking task (i) was actually discrete (a sum of 10 sine waves), there was some question of the accuracy of the spectral density computations. The accuracy of the computations was investigated empirically in several different ways. One way was to implement two known and linear transfer functions in the forward loop of a unity-feedback closed loop on an analog computer. The transfer functions used were $[1+(1/s)]$ and K/s . Sinusoidal frequency responses of the setup verified each transfer function, and then closed-loop data were taken on FM tape for both the sum of 10 sine waves input and a random input to the closed loop. By using the spectral density technique, the closed-loop data were analyzed and compared with the transfer function $[1+(1/s)]$. For the sum of 10 sine waves data, the results were excellent except at the two lowest frequencies of the input.

The computational technique was also verified by having eight data runs analyzed independently for Y_p on the analog spectral analyzer at the Franklin Institute in Philadelphia, Pa. The results of these analyses are shown in figures 12 to 19. The agreement is seen to be very acceptable for most cases.

DATA PRESENTATION.—The primary presentation of the data is in terms of frequency response plots of both magnitude and phase for each configuration. Both Y_p and $Y_p Y_c$ are presented and are shown in figures 20 to 30. All three pilots are represented on each figure by an identifying symbol which is placed at the average value for his data at that point. The lines at each point represent plus and minus the standard deviation of the data and the standard deviation is determined by using the divisor $N-1$ where N is the number of data values available. At each value of frequency, the three points which represent the pilots are spread for clarity, and the frequency at which the points occur is indicated by a tick mark on the frequency scale. The data for the lowest frequency of the input have not been included in these figures because it is the opinion of the author that these points are very much subject to computational error. Also, the phase points for the highest frequency and often the next to highest frequency should be shifted by 360° , since, to conserve space, only one 360° phase interval is shown.

Several characteristics can be noted from the figures. First, the general consistency of the data is very good. During the days in which the flight data were obtained, the air was very stable and almost no turbulence was encountered. The inflight data variability is consistent with that of the ground-based simulator data. Thus comparisons of flight and ground data can be made with assurance.

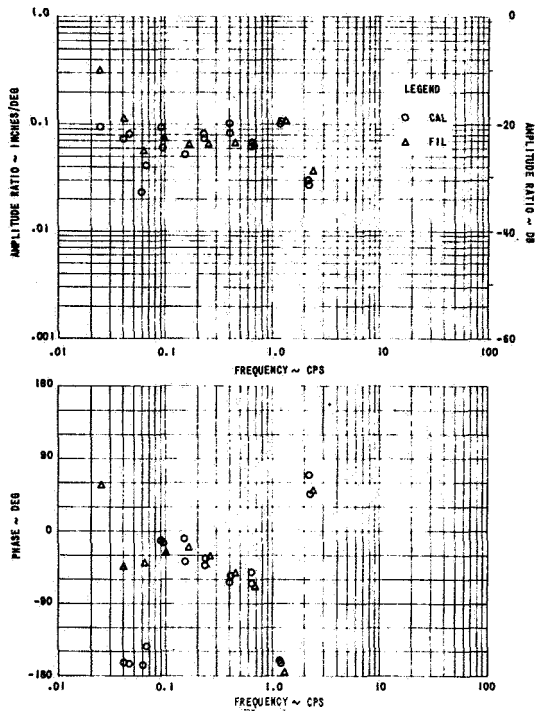


Figure 12.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element K/s of run 25.

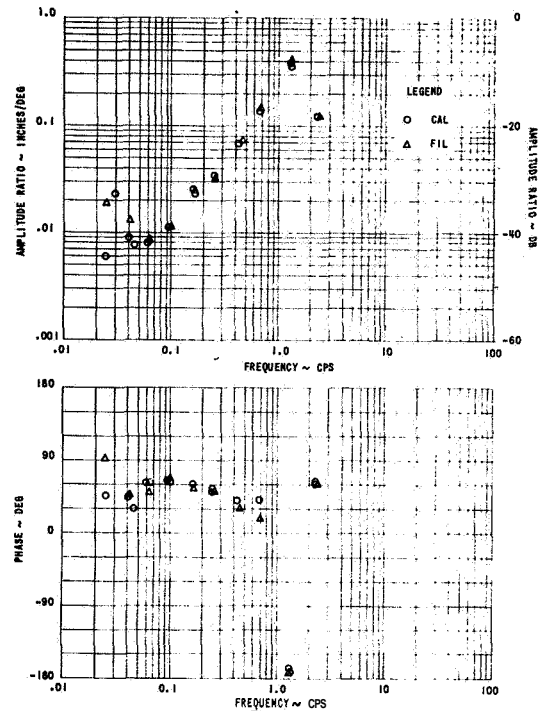


Figure 13.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element $K/s [s-(1/\tau)]$ for run 4.

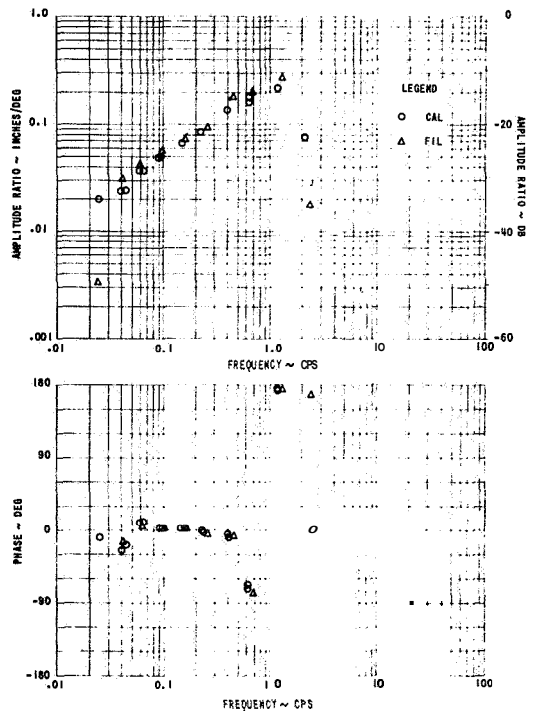


Figure 14.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element A-2 ground for run 39.

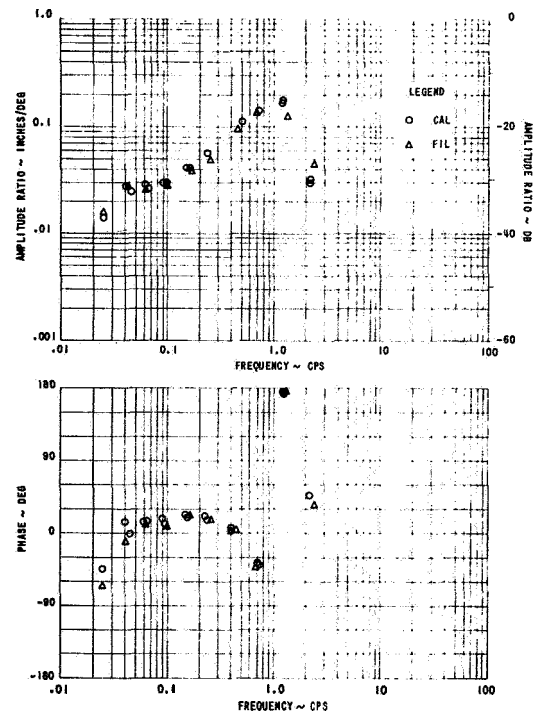


Figure 15.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element A-2*ground for run 69.

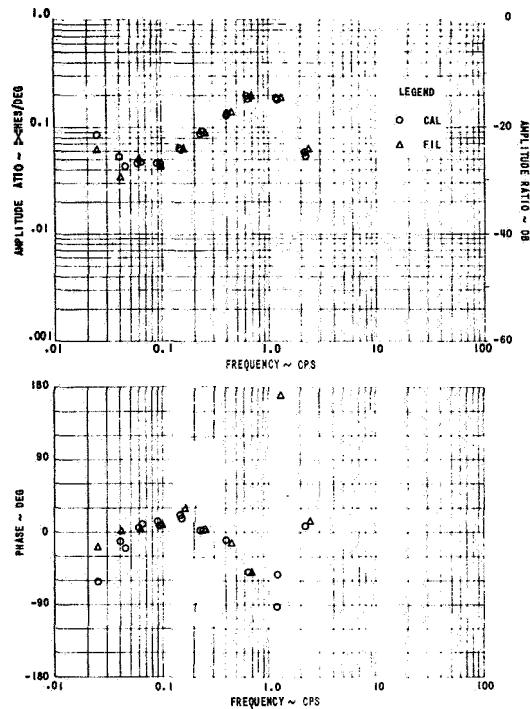


Figure 16.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element A-2*ground for run 60.

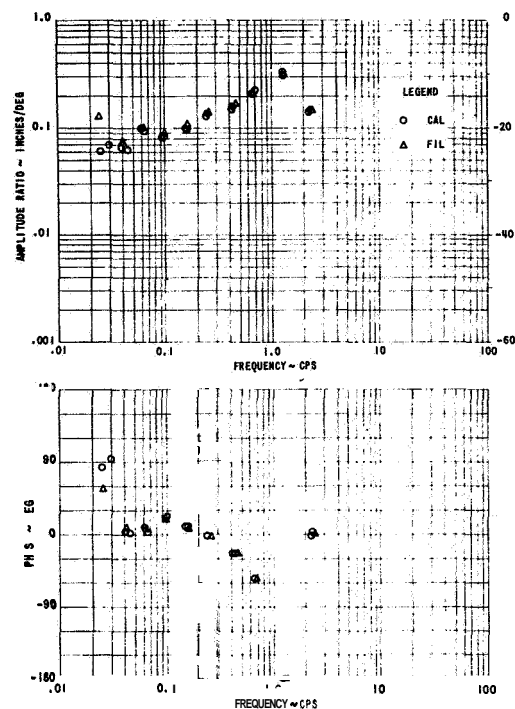


Figure 17.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y with controlled element A-2*flight for run 90.

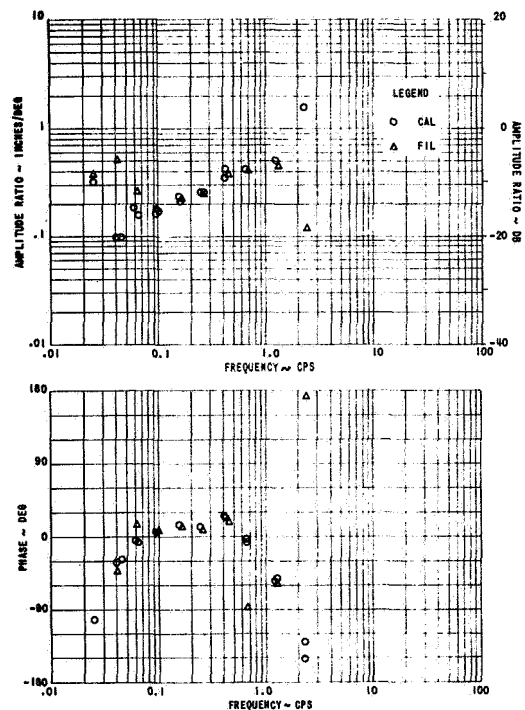


Figure 18.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element A-2 flight for run 91.

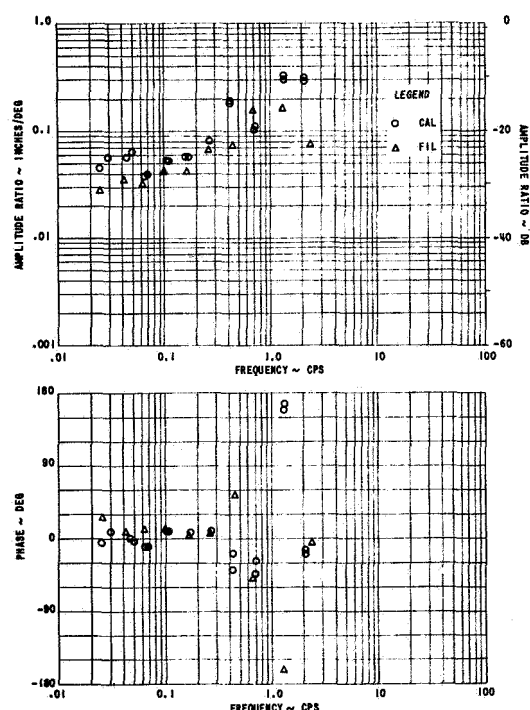


Figure 19.—Comparison of analyses by Cornell Aeronautical Laboratory (CAL) and Franklin Institute (FIL) of pilot transfer characteristic Y_p with controlled element A-2 flight for run 118.

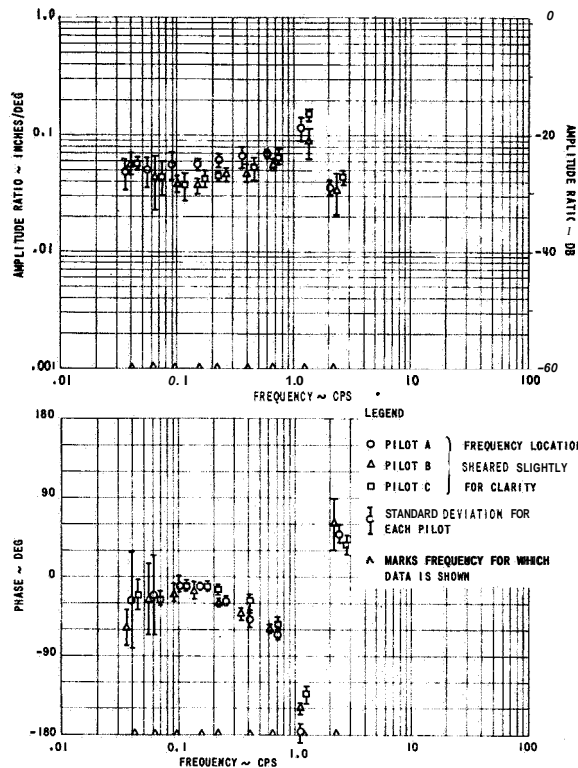


Figure 20.—Pilot transfer characteristic Y_p with controlled element K/s .

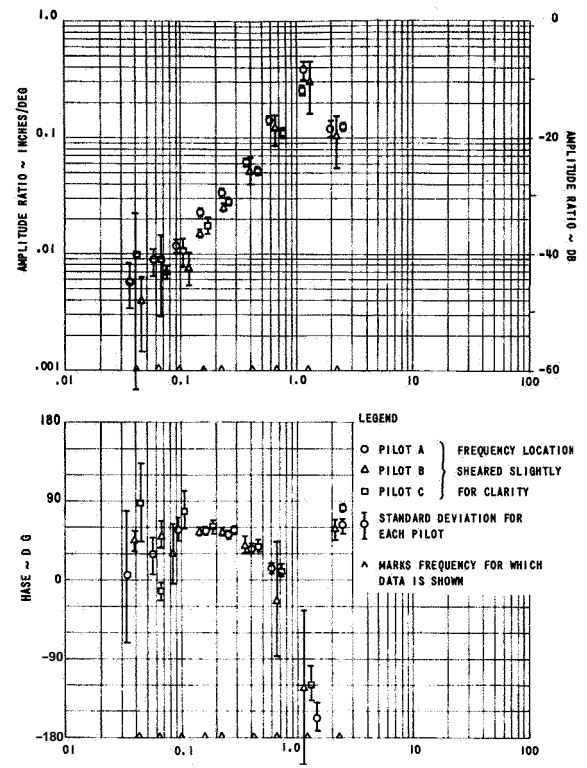


Figure 21.—Pilot transfer characteristic Y_p with controlled element $K/s [s-(1/\tau)]$.

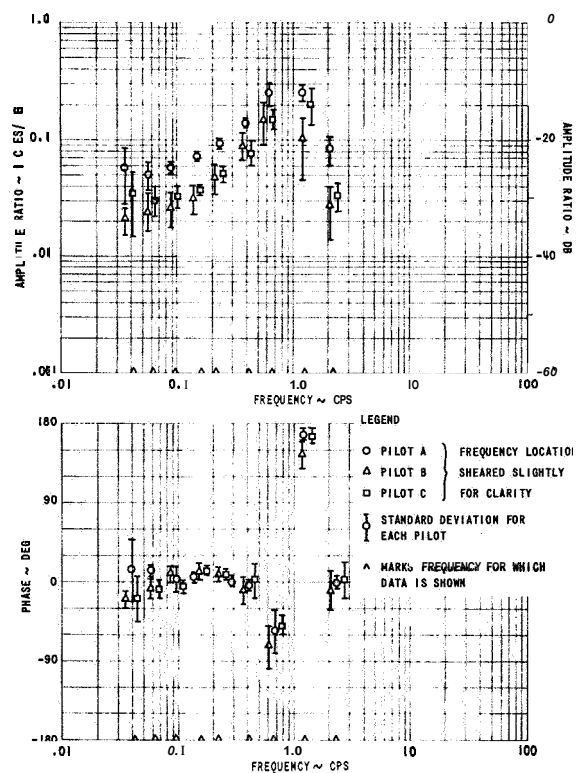


Figure 22.—Pilot transfer characteristic Y_p with controlled element A-2 ground.

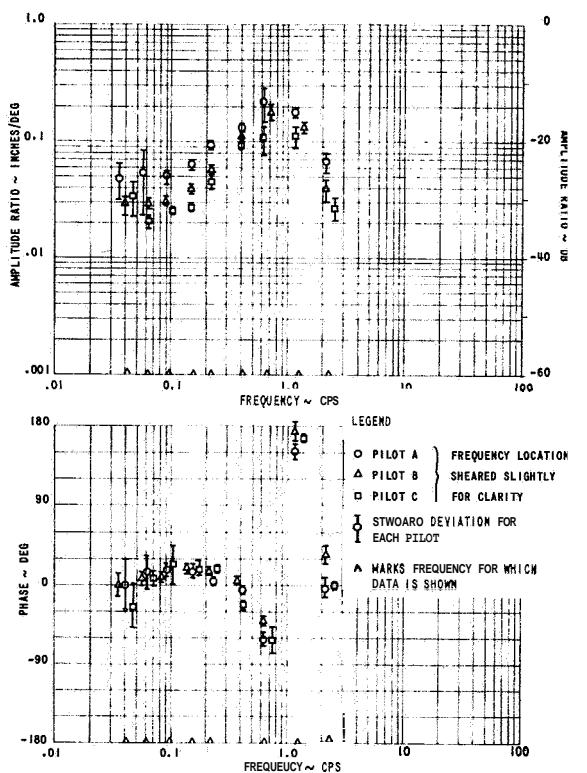


Figure 23.—Pilot transfer characteristic Y_p with controlled element A-2* ground.

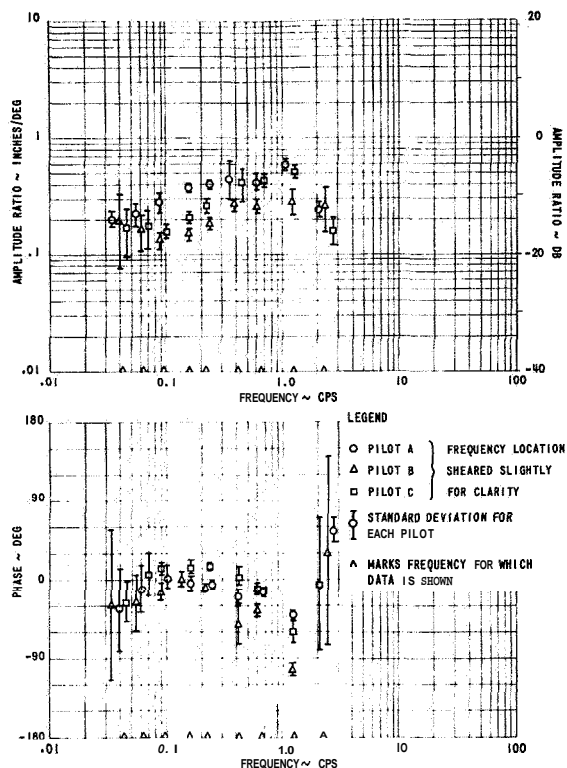


Figure 24.—Pilot transfer characteristic Y_p with controlled element A-2 flight.

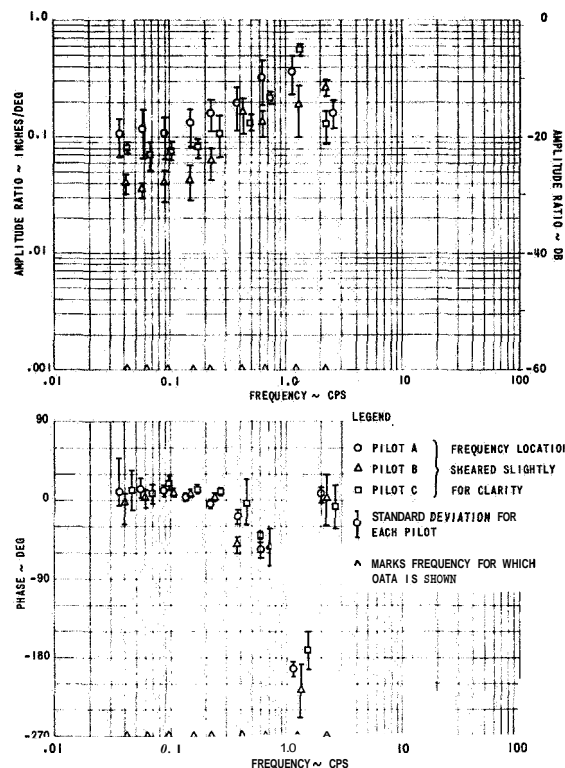


Figure 25.—Pilot transfer characteristic Y_p with controlled element A-2* flight.

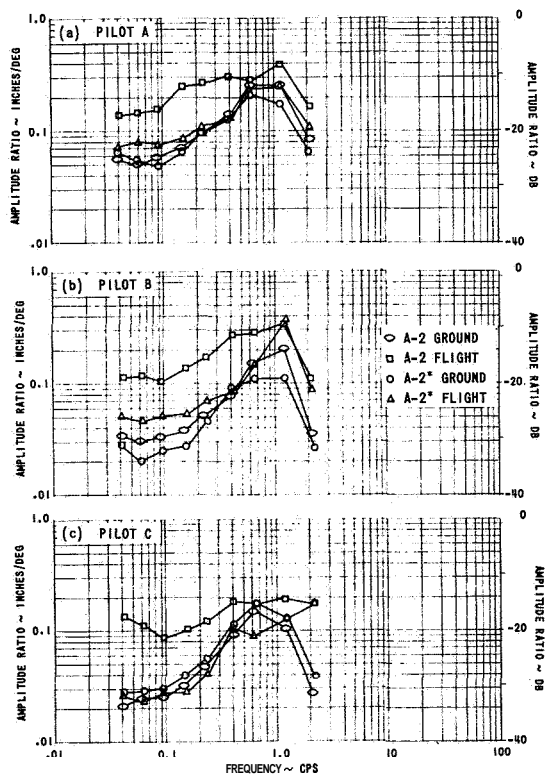


Figure 26.—Comparison of flight and ground pilot transfer characteristics.

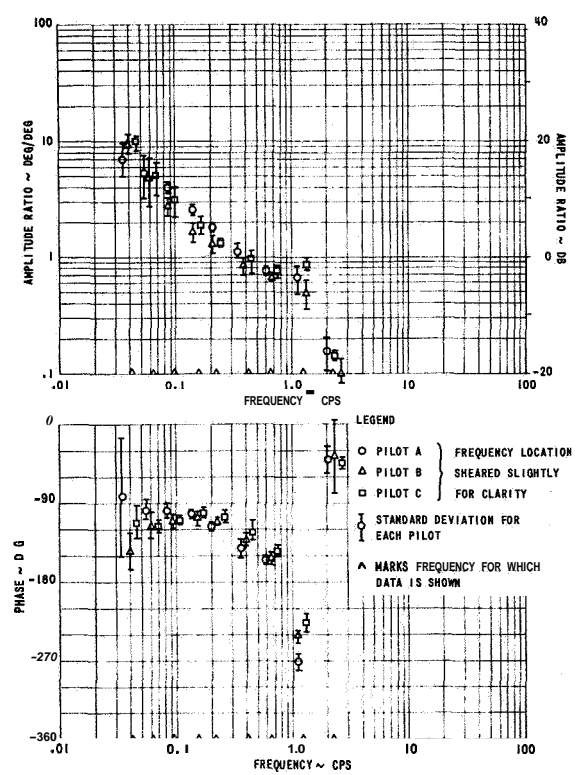


Figure 27.— $Y_p Y_c$ with controlled element K/s .

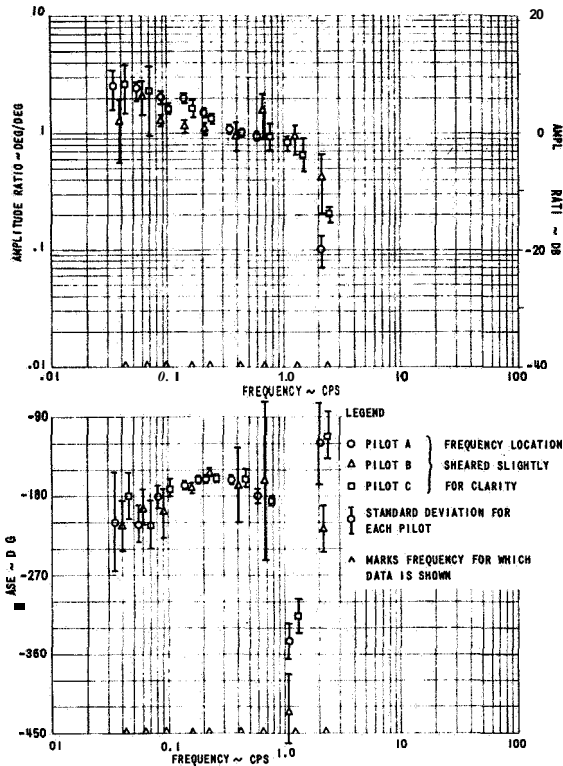


Figure 28.— $Y_p Y_c$ with controlled element $K/s [s-(1/\tau)]$.

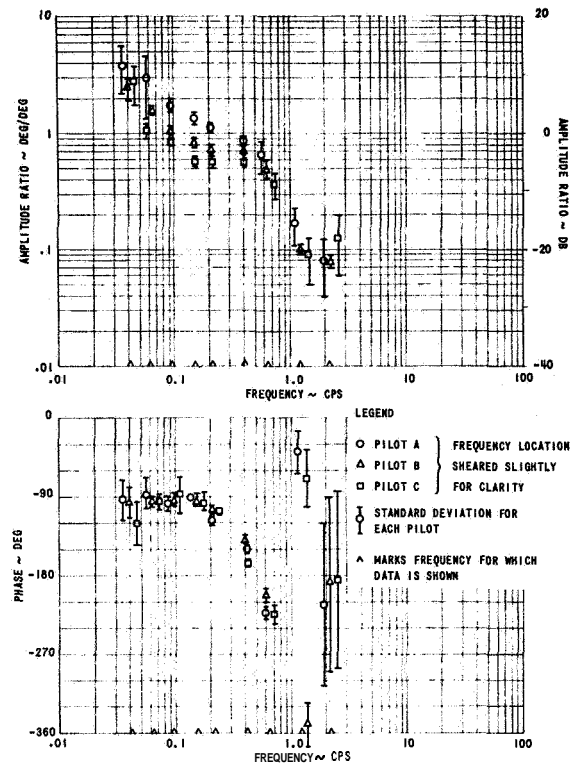


Figure 29.— $Y_p Y_c$ with controlled element A-2* ground.

An interesting characteristic, which is a curious one, is that of the data tending to show noninteger slopes. This tendency is not necessarily inconsistent with previous results since it is suggested in reference 2 that "for the controlled elements with nonzero poles, the open-loop describing function in the region of crossover tends to be somewhat less than -20 dB/decade." However, this tendency seems to be stronger in the data presented herein. The A-2 flight configuration indicates the strongest tendencies for all three pilots to produce the noninteger slope characteristic functions. For the A-2* ground and flight and the A-2 ground configurations, the tendency toward a noninteger slope is not as marked as that for the A-2 flight configuration, and pilot A exhibits the tendency more strongly than do pilots B and C.

From looking at all of the Y_p data, it is easily noted that pilot A always uses a higher gain than the other pilots use.

FLIGHT VERSUS GROUND COMPARISON.—A comparison of the Y_p data for the A-2 and A-2* configurations for both ground and flight is made for each pilot in figures 26(a), 26(b), and 26(c). For these figures, the K_p for the flight data has been adjusted to account for the difference in the original data that is caused by the difference between $K_{cflight}$ and $K_{cground}$. The adjustment is made by multiplying the ratio $K_{cflight}$ per $K_{cground}$ times the flight data.

The most evident implication of the figures is that there is a strong influence by the display on the gain that the pilot chooses. Each pilot substantially increases his gain with the heads-up, visual display over the purely instrument display of either ground or flight. The increase in gain appears to be a function of frequency, but the function is different for each of the three pilots. For pilot A (fig. 26(a)) the gain for heads-up, visual flying is larger than for the instrument flying throughout the frequency range, although the gains for the two cases approach

each other as the frequency increases. The consistency of pilot A's data for the instrument display runs of both ground simulator and actual flight environments may indicate that for this pilot, the effects of nonvisual motion cues are negligible for roll motions between $\pm 8^\circ$ of bank angle. For pilot B, as seen in figure 26(b), there appear to be a major effect of display and a lesser effect of an increase in pilot gain for flight versus ground simulation as shown by contrasting A-2* flight with both A-2 ground and A-2* ground. For this pilot, the difference between instrument flight and visual flight disappears at the two highest frequencies. For pilot C, as seen in figure 26(c), the difference in display is predominant up to a frequency of 0.5 cps and thus the nonvisual motion cues appear to be negligible. The data then, first coalesce and then diverge so that they end at the highest frequency, with both flight configurations being considerably different from both ground-simulation configurations. The complete physical significance of the differences displayed among the pilots is not understood at this writing.

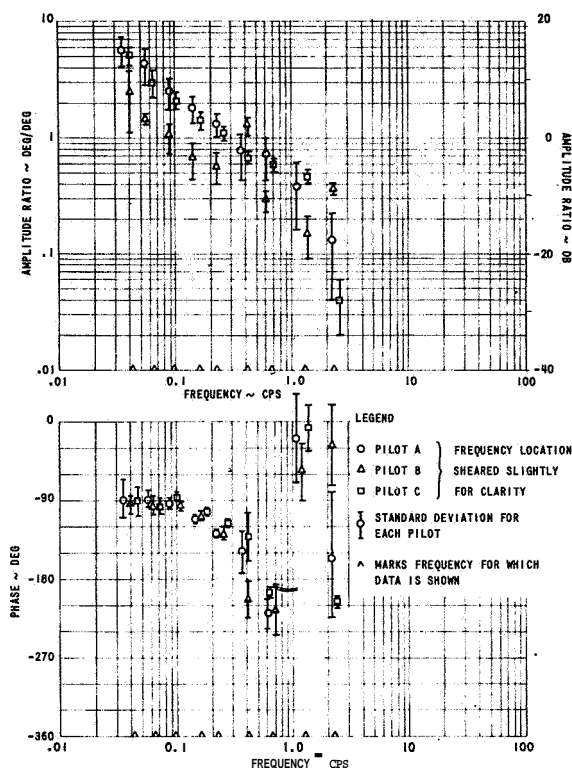


Figure 30, $-Y_p Y_c$ with controlled element A-2* flight.

CROSSOVER FREQUENCY.—The series of transfer characteristics shown in figures 27 to 30 are the magnitude and phase of $Y_p Y_c$.

The data for the simple integrator controlled element as displayed in figure 27 are very similar to figure 41(a) of reference 1, except for a difference in crossover frequencies. The crossover frequency in the referenced data is 0.64 cps, and for the data in figure 27 it is approximately 0.4 cps.

For the divergent configuration, $\frac{K}{s[s-(1/\tau)]}$, data displayed in figure 28, the pilots' data coalesce at the crossover frequency of approximately 0.5 cps, and this tendency is consistent with previous data presented in figure 44(c) of reference 1. However, the data from the reference possess a slope of -1 through the crossover frequency, and the data in figure 28 definitely display a noninteger slope through the crossover frequency.

From figures 29 and 30 for the A-2* configuration for both ground and flight, it is noted that a considerable range in crossover frequency (for zero dB gain) is exhibited among the pilots and that pilot B exhibits a different crossover frequency for flight than he does for ground simulation.

THE STATIONARITY ASSUMPTION.—The assumption that a pilot's behavior is stationary for a 4-minute tracking run has been investigated for 100 seconds of run 106 by using the deterministic theory for measuring time variability that has been developed by Wierwille and which is presented in reference 2. Run 106 is a flight run with configuration A-2 (visual, heads-up) flown by pilot A.

The results of the analysis are presented in figure 31. In the figure, the presentation of the response of the pilot model to a step input has two time axes, t and τ . The time axis t

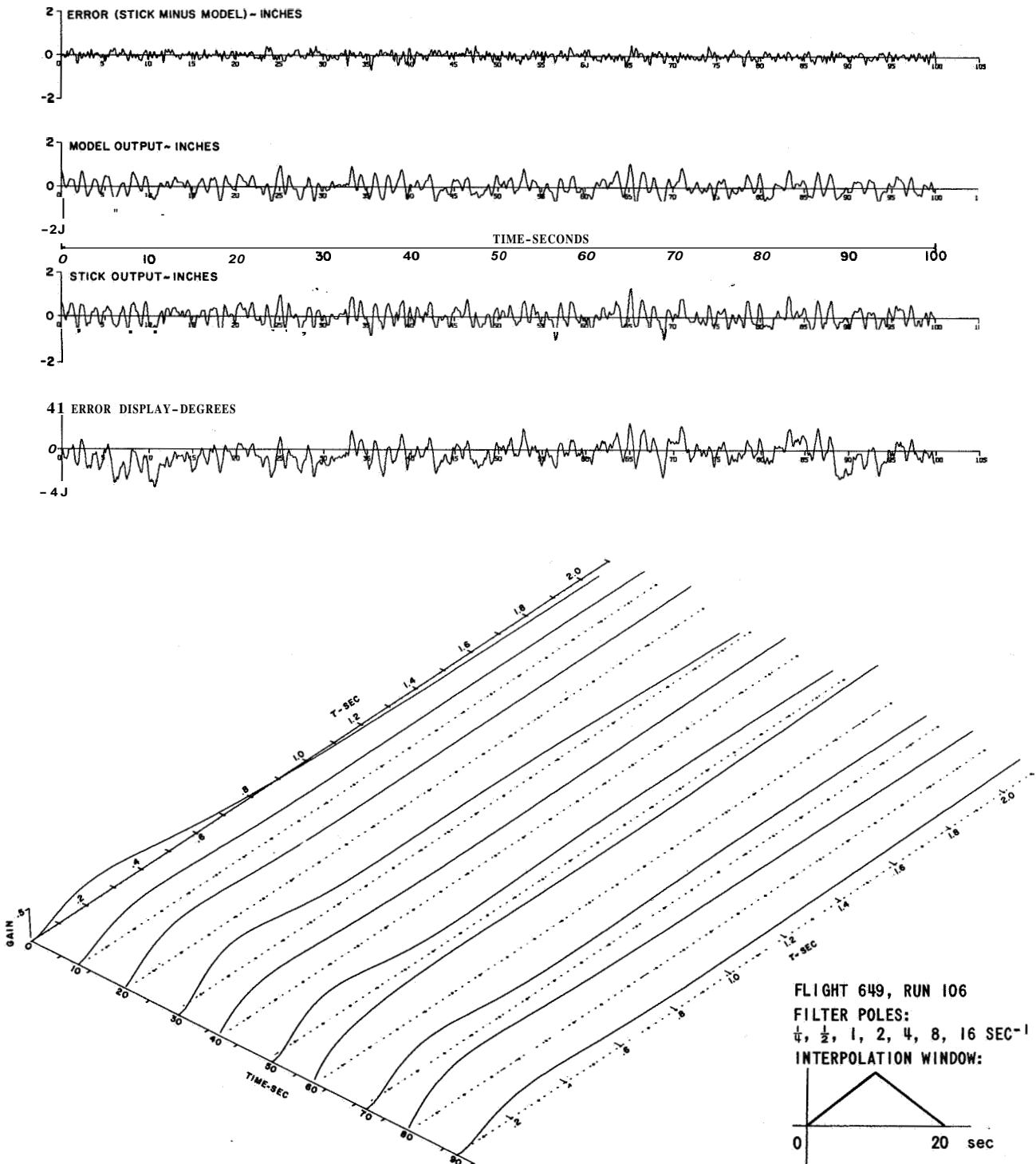


Figure 31.—Output of time-varying pilot model.

represents the time throughout the run at which the pilot model is sampled with a step input. The time axis τ is the axis for plotting the time history of the output of the pilot model at time t . If the pilot model varies from time t_1 to t_2 , then the variation will be depicted by different step-response time histories at these two times. Because the model characterizes the pilot, changes in the model reflect changes in the pilot. The time base used in the truncation

function of the time-varying analysis of this run is 20 seconds, and the total data analyzed is 100 seconds long. The model consists of seven poles and six zeros. The poles ranged from 0.25 to 16 rad/sec.

The first step response shown in the figure appears anomalous and perhaps can be traced to either anomalous behavior of the pilot or to improper initialization of the model. The limited scope of this investigation precluded detailed investigation of this peculiarity.

No definite pilot-reaction time is apparent in the step responses. This unusual result is attributed to the fact that the digital data that were available for analysis were limited to a maximum of eight data points per second and hence a pole of high enough frequency to define the pilots' reaction time could not legitimately be included in the model.

The NISE (normalized, integral square error) figure of merit of 18.39 is a percentage which expresses the fact that the output of the model, for the same input as the pilot had, accounts for 81.61 percent of the variance of the pilots' output. A comparison of the "model output" and the pilot's "stick output" which are included in figure 31 indicates that the wave forms of the two outputs are very similar. A point-by-point comparison of the two outputs is included in the figure as the trace which is identified as "error (stick minus model)" and which is the arithmetic difference between the two output traces. The trace identified as "error display" is the input to the pilot and to the model.

The time varying data presented in figure 31 indicate that pilot A, for the one run analyzed, is rather consistent and therefore it may be concluded that there is little nonstationarity exhibited in the range of periods from 10 to 90 seconds. This conclusion is based on considering all of the step responses except the first one.

LIMITED VARIANCE ANALYSIS.—Some analyses of variance have been performed on the amplitude data for Y_p and $Y_p Y_c$ for the ground simulator runs. The results of these analyses are presented in table 4. In the table, N. S. means not significant, two asterisks mean significant at the 1-percent level, and three asterisks mean significant at a level very much smaller than the 1-percent level. The analysis is based upon taking the first three runs performed by each pilot. This approach is used to make whatever effect of learning there might be as consistent as possible among the three pilots and to perform an analysis which does not tend to bias the results toward pilot A who usually performed more runs per configuration than either of the other pilots. The analysis is designed to determine the significance of the exhibited differences in the means of the configurations, the means of the pilots, and the interaction between the pilots and the configurations. Thus the model for the analysis is

$$y_{ijk} = m + t_i + P_j + e_{ijk}$$

where

y a data point

m general mean

t_i treatment i (i=configurations K/s, $\frac{K}{s - \frac{1}{\tau}}$, A-2, A-2*)

P_j pilot j (j=A, B, or C)

e_{ijk} error associated with point ijk

k runnumber

TABLE 4.—ANALYSIS OF VARIANCE RESULTS

| | | Frequency | | | | |
|----------------|-----------|------------|-----------|-----------|-----------|----------|
| | | 0.0625 cps | 0.237 cps | 0.404 cps | 0.642 cps | 1.20 cps |
| Configurations | Y_P | (a) | (a) | (a) | | (a) |
| | $Y_p Y_c$ | (a) | (a) | (a) | (a) | |
| Pilots | Y_P | (b) | (a) | (a) | | (a) |
| | $Y_p Y_c$ | (c) | (a) | (a) | (c) | |
| Interaction | Y_P | (c) | (a) | (a) | | (b) |
| | $Y_p Y_c$ | (c) | (b) | (b) | (b) | |

^aSignificant at a level very much smaller than 1-percent level

^bSignificant at the 1-percent level

^cNot significant

The analysis was performed at four frequencies, 0.0625, 0.237, 0.4042, and 1.204 cps for the Y_p data to determine if significant differences vary with frequency, and at 0.0625, 0.237, 0.4042, and 0.642 cps for the $Y_p Y_c$ data.

For the variance analysis of the $Y_p Y_c$ data, the values of $Y_p Y_c$ for the A-2 configuration were generated by multiplying the Y_p determined for A-2 with the Y_c determined from the A-2* configuration. This is a legitimate procedure because the only difference between the two configurations is the point of injection of the forcing function.

The most important characteristic displayed in table 4 is the nonsignificant difference among the pilots at a frequency of 0.642 cps, which shows that the three pilots do not differ from one another at this frequency.

CONCLUSIONS AND RECOMMENDATIONS

From the results presented in this paper, the following conclusions and recommendations may be drawn:

(1) It is possible to gather accurate, meaningful, inflight data which describe the dynamic behavior of the human pilot in single-control, compensatory tracking tasks.

(2) The pilot transfer characteristics gathered in the small-disturbance bank-angle tracking task indicate that the pilot behavior in flight is basically similar to his performance in the ground simulator for the same instrument displayed information.

(3) The pilot transfer characteristics gathered in the small-disturbance bank-angle tracking task indicate that pilot behavior in flight is different when the information used by the pilot is obtained from visual reference to the outside world as compared with flight by reference to an instrument display. The main difference is that each pilot uses higher gain when he refers to the outside world than he uses when he refers to instruments. There are also some differences in the dynamic adaptations, and these appear to be somewhat individualistic with the pilot.

(4) Limited examination of the data indicates that for the "good" airplane controlled element, the pilot exhibits relatively constant dynamic behavior over the 4-minute tracking run. Therefore, the stationarity assumption of the data analysis techniques appears to be substantiated.

(5) Often times at the lowest and the highest frequencies of the forcing function input, there appears to be insufficient power to measure accurately the pilot transfer characteristic for the easily flown controlled elements. At the other frequencies of the forcing function, consistently satisfactory data were extracted.

(6) The variance of the extracted transfer characteristics was approximately unchanged between ground and flight tracking runs.

(7) Each of the three pilots exhibited basically similar adaptations to each of the controlled elements. Detailed differences were exhibited in the gain adaptation, particularly that shown by pilot A, who used a consistently higher gain.

(8) The similarity exhibited among the pilots in their adaptation to each controlled element strengthens the potential of correlating measured transfer characteristics of the pilot with his verbal assessments and ratings of handling qualities. The transfer characteristics, if taken for tasks which bear important relationships to typical missions, will provide substantial analytic information, which presently does not exist, about the pilot-vehicle system. This information combined with pilot evaluation comments will significantly improve the understanding of the pilot-vehicle system. Caution is recommended in the interpretation of this conclusion because meaningful, mission-oriented tasks usually involve multivariable, multicontrol roles for the pilot. Appreciable research effort must be expended in both the theoretical and experimental areas of the multivariable, multicontrol problem to extend the techniques of this program to the broad area of handling qualities research.

REFERENCES

1. McRuer, D. T.; Graham, D.; Krendel, E.; Reisener, W. C.: Human Pilot Dynamics in Compensatory Systems—Theory, Models, and Experiments with Controlled Element and Forcing Function Variations. AFFDL-TR-65-15, July 1965.
2. Wierwille, W. W. and Gagne, G. A.: A Theory for the Optimal Deterministic Characterization of the Time-Varying Dynamics of the Human Operator. NASA CR-170, Feb. 1965.

8. Describing Functions for Compensatory Tracking of Sine Waves Plus Noise

A. C. Deure and A. Kuhn
Westinghouse Defense and Space Center

768

15900

A study of compensatory tracking was performed on an analog computer simulation to test the application of the superposition theorem to human tracking performance. Four subjects participated in a $4 \times 4 \times 4 \times 3 \times 2$ analysis of variance design in which the input variables of frequency, amplitude, stick and scope sensitivity, and noise conditions were varied systematically. The performance data of error and stick movement and the ratio of these two measures, that is, the gain, were subjected to an analysis of variance performed on a digital computer.

The results showed that the linearity assumption is not a valid assumption. The variable underlying performance is the average rate of stick motion. Using the rate variable, a transfer function was derived which provides an adequate fit between the empirically derived and the theoretically calculated data.

The conclusions show that error increases and gain decreases as a direct function of average rate of stick motion and that the presence of noise has an effect similar to that of increasing the rate.

Since the latter days of World War II engineers have been designing radar-directed fire control systems in which the human operator closes the loop between the element to be controlled and the sensing device. This design effort requires an adequate match between the characteristics of the control system and those of the human operator. In order to obtain this match, various attempts have been made to derive transfer functions that would describe the human. Such a function still remains a prerequisite for a complete analysis of the entire system, including the human operator. Moreover, with the availability of such a function, a system could be developed which fully utilizes the human operator's capabilities without placing demands on him which exceed his limitations. This report represents an experimental investigation and analysis aimed at providing such a transfer function for the human operator.

Much of the past work on developing transfer functions for the human operator has been summarized by Bekey (ref. 1) and by McRuer et al. (ref. 2), among others. An analysis of this literature reveals that none of these studies have been conducted in accordance with the accepted procedures of experimental design that must be applied in conducting complex experiments. The literature shows that, generally, the following assumptions have been made:

(1) An analysis based upon a random input will provide an adequate vehicle for analyzing the results of a single sinusoidal input (ref. 3).

(2) The human operator's performance can be considered linear. These assumptions have been made so that the mathematical procedures applicable to linear transfer function theory can be used. Such application has persisted in spite of the general admission that the human operator is essentially nonlinear (ref. 4) and that his representation by a linear function is inadequate.

In the development of these linear transfer functions, ad hoc hypotheses were made. The most frequent is the intermittency hypothesis and its related model, the sampled data hypothesis (ref. 1). These assume that the operator samples the input at specific intervals

and responds to it intermittently. Thus, a finite amount of time is assumed to occur between responses even though the task appears to be performed continuously. Recently there have been attempts to improve this representation by postulating a more complicated neuromuscular reaction time function (ref. 5). It is surprising to find these assumptions when Gibbs (ref. 6) has shown that the reaction time hypothesis is untenable by his demonstration that the performance of a highly developed skill, such as tracking, is based on a feedback of stimuli that are learned as control cues. Such cues must necessarily precede the visual recognition of error if the task is to be accomplished with a minimum of error. Conklin (ref. 7) has also shown in a study of tracking which employed different filters that the human is capable of behaving as a continuous error correcting servo. Finally, Brazier has marshalled impressive evidence to show that the nature of neural activity acts in a probabilistic rather than a deterministic fashion so that a single reaction time value could not possibly serve all synapses (ref. 8).

From the foregoing analysis it would appear that a careful examination of the linearity hypothesis and its experimental implications would be of value in clarifying the theoretical problem and in determining what factors have not been considered.

The linearity hypothesis, or the super-position theorem, states that a system will respond to changes in the frequency of the input and that such a response will be essentially independent of the amplitude of the input. When these statements are translated into mathematical terms for the linear transfer function model the following conditions result:

Let

e_o measure of the output

e_i measure of the input.

If

$$e_o = A_o (\cos \omega t + j \sin \omega t)$$

and

$$e_i = A_i (\cos \omega t + j \sin \omega t)$$

then $A_o = A_i$ if the ratio of output to input is independent of amplitude and the ratio would be dependent upon frequency alone. Therefore, any change of input amplitude should not yield any change in the value of the ratio for a constant input frequency. Assuming that this is in

fact the case, gain, which is the ratio of output to input, $\frac{e_o}{e_i}$, should remain constant over a given range of amplitudes, that is,

$$\frac{e_{o1}}{e_{i1}} = \frac{e_{o2}}{e_{i2}} = \frac{e_{o3}}{e_{i3}}$$

If this relationship holds true, the expression $G = \frac{e_o}{e_i}$ can be written as $G e_i = e_o$; that is, output

is proportional to input. Therefore, if the output is independent of amplitude, a plot of gain as a function of amplitude for a particular frequency should yield a straight line with zero slope. Since the error visually seen on the display is considered the input, the linear

hypothesis requires that the output or control movement change proportionally *so* that gain will in fact remain constant for a given frequency. It was the purpose of the research reported here to examine the assumptions of the linearity hypothesis and to present an alternate approach to the analysis of human tracking behavior if these linearity assumptions can be shown not to hold. This alternate approach is based upon the early unpublished research of A. Kahn which indicated that the average rate of stick motion is a primary variable in tracking. The study adheres to the procedures and techniques of rigid experimental design and control outlined by Fisher (ref. 9).

METHOD

APPARATUS.—The study was performed on an analog computer and a sine wave generator. The display consisted of an oscilloscope on which the target moved in the horizontal dimension. The control was a metal stick, mounted on the arm of the subject's chair, and a tape marked the place at which subjects were to hold the stick. A block diagram of the complete control loop is given in figure 1. A six-channel recorder provided a trace of the following measures for each run:

- (1) Forcing function
- (2) Instantaneous stick position
- (3) Instantaneous error
- (4) Integrated error**
- (5) Integrated stick
- (6) Instantaneous dot position.

Only channels (4) and (5) were used for the purpose of the present analysis.

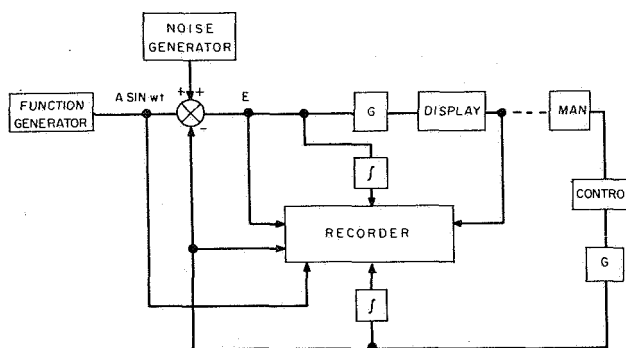


Figure 1.—Block diagram of apparatus.

PROCEDURE.—Four male operators were used, ranging in age between 23 and 39; all had 20-20 vision. They tracked in sessions of 1 or 2 hours with rest periods after about every tenth trial. The viewing distance from subject to scope face was 28 inches. Each observer was given the following instructions:

- (1) Center the stick and hold it by the taped section when the "ready" signal is given.
- (2) Start tracking at the "go" signal by manipulating the stick to return the dot to the center position and keep it centered for the whole of each 30-second run.

When the operator indicated that he understood the compensatory tracking task, about five trial runs were made. The experimenter programmed frequency, amplitude, stick sensitivity, and scope sensitivity controlled by potentiometers, gave the "ready" signal, turned on the recorder, and when the record showed that the signal was going through zero, said "go" and turned on the computer. The subject then tracked the signal until the computer clicked off, which was automatically timed to occur after 30 seconds and which returned the signal to zero position. A new set of input variables was then programmed, and the next run started.

EXPERIMENTAL VARIABLES AND DESIGN.—The experimental variables are given in table 1. The forcing function is a sine wave whose four frequencies are given in the first row of table 1. The oscillator amplitude values are the amplitudes appearing on the scope with scope sensitivity 1 (i.e., 5 deg/cm) when no stick action is applied. These oscillator amplitudes can be translated into stick amplitudes by multiplying them by the stick sensitivities (ignoring the scope sensitivity since the oscillator and stick signals are summed directly).

TABLE 1.—EXPERIMENTAL VARIABLES

| | Variable number | | | |
|------------------------|-----------------|------------|------------|----------|
| | 1 | 2 | | 4 |
| Frequency | 0.03 cps | 0.06 cps | 0.10 cps | 0.50 cps |
| Oscillator
amp, P-P | 2 cm | 4 cm | 6 cm | 8 cm |
| Stick
sensitivity | 5 deg/cm | 5 deg/2 cm | 5 deg/3 cm | |
| Scope
sensitivity | 5 deg/cm | 5 deg/2 cm | 5 deg/5 cm | |
| Noise | Present | Absent | | |

Stick sensitivity refers to the fact that variable gain can be placed on dot displacement on the scope per degree of stick movement. Scope sensitivity is the a priori gain in dot displacement on the scope by the multiplicative factors of 1, 2, or 5. The presence of noise refers to Gaussian noise, 0 to 27 cps, filtered by simple lag at 1 cps.

The variables shown in table 1 with four observers combine into a $4 \times 4 \times 3 \times 3 \times 2$ analysis of variance design, yielding 288 conditions, or runs, for each observer. This means that every variable occurs in combination with every other variable. For example, frequency 1 is paired once with each amplitude, once with each stick sensitivity, once with each scope sensitivity, and once with each noise-on and noise-off, such as the following example: frequency 1, amplitude 2, stick 2, scope 3, and noise. All combinations were run once with noise and once without noise. Each run was automatically timed (and recorded) for 30 seconds, and the 288 combinations of experimental conditions were arranged into a different random order of presentation for each observer.

As shown below, the calculation of the average rate of stick motion does not involve the variable of scope sensitivity. However, since the work of Kahn and Mazina (ref. 10) showed this variable to be critical in display design, it was included here in order to test its effect on operator performance.

RATE.—By combining frequency, amplitude, and stick sensitivity, 30 average rates of stick motion were derived by the following equation:

$$\frac{S_{ST}(4A)}{T_F} = R_{ST} \quad (1)$$

where

S_{ST} stick sensitivity

A amplitude of sine wave

T_F time period of one cycle of a given frequency

R_{ST} average rate of stick motion

Thus (the term 4A refers to four times the amplitude, i.e., the completely generated sine wave), we know for each combination of experimental variables which average rate of stick

motion the input represents. For the purpose of this presentation, all these calculated rates were converted into radians per second.

Before discussing our findings, a word of explanation on the difference between the terms stick movement and average rate of stick motion: Stick movement is a subject performance measure and refers to the direct measure taken off channel 5 of the tracking record. As with the error measure, the trace's absolute deviation from zero is summed and integrated directly by the analog computer. The final integral is measured off the tape and divided by 30, giving an absolute angular measure. The data here always represent the mean measures for all four subjects. Average rate of stick motion is an input variable to which the tracker's hand responds continually. As mentioned earlier, it is made up of frequency, amplitude, and stick sensitivity. It represents the demand made on the operator at any point in time, and is written in terms of angular unit per unit of time.

RESULTS AND DISCUSSION

The information from channels 4 and 5, that is, integrated error and integrated stick movement, was measured and averaged for each subject during each run. These measures were then subjected to an analysis of variance performed on a digital computer. Three analyses were obtained: one for error, one for stick movement, and one for the ratio of stick movement error (i.e., output/input)=gain. Each of these analyses indicates the influence which each variable or combination of variables exerts on the performance measure and assesses the statistical significance of this influence. Table 2 presents the summary of the analysis of variance for the error measure; table 3 presents a summary for stick movement; and table 4 presents a summary for the gain ratio. The probabilities used for determining levels of significance are 0.01 and 0.05, and only those sources of variance that are significant at these levels were tabulated and are discussed here. The interactions due to subjects will not be discussed because no effort was made to control this variable. Although scope sensitivity appears as a significant variable in tables 2 and 4 and stick sensitivity is shown to be significant in all three analyses, the difference between performance measures as a function of either the scope or the stick sensitivities used was not large enough to present the data from each of the nine combinations of these variables. We therefore chose, for all figures presented here, only data reflecting the combination of stick sensitivity 2 with scope sensitivity 2 as being fully representative of the families of curves obtained with all the other stick and scope sensitivity combinations.

An examination of table 2 indicates that those variables which contribute to the determination of rate at the hand, that is, frequency, amplitude, and stick sensitivity, are significant. The table further shows that noise was a significant source of variance, while in table 3, which evaluates stick movement, noise was not a significant factor. These two findings suggest that noise is a source of variance in the error data because the operator cannot determine whether the error perceived on the scope as a visual input is due to noise or to actual error. In moving his control, however, the operator tends to ignore the noise and attempts to approximate the overall movement. The importance of average rate of stick motion as a source of variance, as compared to noise, is easily seen by comparing the sum of squares for noise to that for frequency amplitude and stick sensitivity each. Table 3 shows essentially the same information, and it demonstrates that the visual input in terms of scope sensitivity has no effect on the amount of stick movement. Similarly, the importance of average rate is demonstrated by the large sum of squares due to amplitude and stick sensitivity and the higher interactions developed by combinations of variables that yielded rate of stick motion. Table 4 shows the effect of each of these two previous analyses combined so that gain is influenced by the variables that influence error and by those that influence stick movement. To the degree

TABLE 2.—PARTIAL SUMMARY OF ANALYSIS OF VARIANCE FOR ERROR

| Source | Sum of squares | df | Variance estimate | F-ratio |
|--|----------------|------|-------------------|----------------------|
| Subjects | 33.2283 | 3 | 11.0761 | ^a 84.81 |
| Noise | 8.2356 | 1 | 8.2356 | ^a 63.06 |
| Scope sensitivity | 1.3622 | 2 | .6811 | ^a 5.21 |
| Frequency | 618.0917 | 3 | 206.0306 | ^a 1577.57 |
| Amplitude | 98.9031 | 3 | 32.9677 | ^a 252.43 |
| Stick sensitivity | 201.1628 | 2 | 100.5814 | ^a 770.15 |
| Subjects × scope sensitivity | 2.1833 | 6 | .3639 | ^b 2.78 |
| Subjects × frequency | 48.6148 | 9 | 5.4016 | ^a 41.36 |
| Subjects × amplitude | 10.5479 | 9 | 1.1720 | ^a 8.97 |
| Subjects × stick sensitivity | 14.9583 | 6 | 2.4930 | ^a 19.09 |
| Noise × frequency | 1.2469 | 3 | .4156 | ^b 3.18 |
| Noise × stick sensitivity | 1.3974 | 2 | .6987 | ^a 5.35 |
| Frequency × amplitude | 111.1041 | 9 | 12.3449 | ^a 94.52 |
| Frequency × stick sensitivity | 135.9438 | 6 | 22.6573 | ^a 173.48 |
| Amplitude × stick sensitivity | 34.3389 | 6 | 5.7231 | ^a 43.82 |
| Subjects × frequency × amplitude | 14.6185 | 27 | .5414 | ^a 4.14 |
| Subjects × frequency × stick sensitivity | 30.7616 | 18 | 1.7090 | ^a 13.08 |
| Subjects × amplitude × stick sensitivity | 8.9382 | 18 | .4966 | ^a 3.80 |
| Noise × frequency × stick sensitivity | 3.3502 | 6 | .5584 | ^a 4.27 |
| Frequency × amplitude × stick sensitivity | 34.1359 | 18 | 1.8964 | ^a 14.52 |
| Subjects × frequency × amplitude × stick sensitivity | 14.4893 | 54 | .2683 | ^a 2.05 |
| Residual | 14.1012 | 108 | .1306 | |
| Total | 1538.1851 | 1151 | | |

^aP=0.01^bP=0.05

that there is a communality of influence, these interactions suggest that a basic variable exists. We suggest that this variable is represented by the concept of average rate of stick motion.

TABLE 3.—PARTIAL SUMMARY OF ANALYSIS OF VARIANCE FOR STICK MOVEMENT

| Source | Sum of squares | df | Variance estimate | F-ratio |
|--|----------------|------|-------------------|----------------------|
| Subjects | 60.1120 | 3 | 20.0373 | ^a 7.94 |
| Frequency | 213.4796 | 3 | 71.1599 | ^a 28.19 |
| Amplitude | 2962.2296 | 3 | 987.4099 | ^a 3912.08 |
| Stick sensitivity | 3904.9787 | 2 | 1952.4894 | ^a 7735.69 |
| Subjects x frequency | 139.2637 | 9 | 15.4737 | ^a 61.31 |
| Subjects x amplitude | 9.3615 | 9 | 1.0402 | ^a 4.12 |
| Subjects x stick sensitivity | 26.9514 | 6 | 4.4919 | ^a 17.80 |
| Noise x frequency | 2.1382 | 3 | .7127 | ^b 2.82 |
| Frequency x amplitude | 63.3507 | 9 | 7.0390 | ^a 27.89 |
| Frequency x stick sensitivity | 121.1862 | 6 | 20.1977 | ^a 80.02 |
| Amplitude x stick sensitivity | 653.7165 | 6 | 108.9528 | ^a 431.67 |
| Subjects x frequency x amplitude | 47.7682 | 27 | 1.7692 | ^a 7.01 |
| Subjects x frequency x stick sensitivity | 53.1267 | 18 | 2.9515 | ^a 11.69 |
| Subjects x amplitude x stick sensitivity | 9.5738 | 18 | .5319 | ^b 2.11 |
| Frequency x amplitude x stick sensitivity | 32.0994 | 18 | 1.7833 | ^a 7.07 |
| Subjects x frequency x amplitude x stick sensitivity | 28.1749 | 54 | .5218 | ^a 2.07 |
| Residual | 27.2564 | 108 | .2524 | |
| Total | 8562.7063 | 1151 | | |

^aP=0.01^bP=0.05

The interaction terms in table 4 further indicate that gain is not invariant with amplitude for a given frequency. This is demonstrated in figure 2 which plots error, stick movement, and gain as functions of amplitude and frequency. In this figure, the gain curves show a change with amplitude and frequency for both the noise and no-noise conditions. It approaches linearity only in those instances where the corresponding error curve rises steeply. Thus, a comparison of the error curve for frequency 4 (0.50 cps) with frequency 1 (0.03 cps) shows that the gain changes markedly for frequency 1 while the error changes very little. On the other hand, the gain for frequency 4 is practically constant while the error increases markedly. Figure 2 also shows that there was very little difference in stick movement as a function of frequency in both the noise and no-noise conditions. Error, however, does increase in the noise condition while the amount of stick movement remains the same. As a

TABLE 4.—PARTIAL SUMMARY OF ANALYSIS OF VARIANCE FOR GAIN

| Source | Sum of squares | df | Variance estimate | F-ratio |
|---|----------------|----|-------------------|----------------------|
| Subjects | 2532.9722 | 3 | 844.3241 | ^a 109.80 |
| Noise | 4333.7569 | 1 | 4333.7569 | ^a 562.80 |
| Scope sensitivity | 565.6302 | 2 | 282.8151 | ^a 36.70 |
| Frequency | 29223.7830 | 3 | 9741.2610 | ^t 1265.00 |
| Amplitude | 4193.0885 | 3 | 1397.6962 | ^a 181.39 |
| Stick sensitivity | 754.3585 | 2 | 377.1793 | ^a 48.90 |
| Subjects x noise | 377.1511 | 3 | 125.7170 | ^a 15.00 |
| Subjects x frequency | 867.4302 | 9 | 96.3811 | ^a 12.56 |
| Subjects x amplitude | 611.9496 | 9 | 67.9944 | ^a 8.86 |
| Noise x scope sensitivity | 166.1059 | 2 | 83.0530 | ^a 10.82 |
| Noise x frequency | 2265.4492 | 3 | 755.1497 | ^a 98.42 |
| Noise x amplitude | 121.3577 | 3 | 40.4526 | ^a 5.27 |
| Noise x stick sensitivity | 800.2691 | 2 | 400.1346 | ^a 52.15 |
| Scope sensitivity x frequency | 368.8384 | 6 | 61.4731 | ^a 8.01 |
| Scope sensitivity x stick sensitivity | 221.0946 | 4 | 55.2737 | ^a 7.20 |
| Frequency x amplitude | 3222.6578 | 9 | 358.0731 | ^a 46.67 |
| Frequency x stick sensitivity | 900.8097 | 6 | 150.1350 | ^a 19.57 |
| Amplitude x stick sensitivity | 130.3950 | 6 | 21.7325 | ^b 2.83 |
| Subjects x noise x frequency | 136.8521 | 9 | 15.2058 | ^b 1.98 |
| Subjects x noise x amplitude | 252.0789 | 9 | 28.0088 | ^a 3.65 |
| Subjects x noise x stick sensitivity | 154.4852 | 6 | 25.7475 | ^a 3.35 |
| Subjects x frequency x amplitude | 777.0572 | 27 | 28.7799 | ^a 3.75 |
| Subjects x frequency x stick sensitivity | 236.2586 | 18 | 13.1255 | ^b 1.71 |
| Noise x scope sensitivity x frequency | 141.1318 | 6 | 23.5220 | ^a 3.06 |
| Noise x scope sensitivity x stick sensitivity | 60.9809 | 4 | 15.2452 | ^a 19.87 |
| Noise x frequency x amplitude | 323.0573 | 9 | 35.8953 | ^a 4.68 |
| Noise x frequency x stick sensitivity | 1397.0956 | 6 | 232.8493 | ^a 30.35 |
| Noise x amplitude x stick sensitivity | 351.6787 | 6 | 58.6131 | ^a 7.64 |
| Frequency x amplitude x stick sensitivity | 240.2005 | 18 | 13.3445 | ^b 1.74 |

TABLE 4.—PARTIAL SUMMARY OF ANALYSIS OF VARIANCE FOR GAIN—Concluded

| Source | Sum of squares | df | Variance estimate | F-ratio |
|---|----------------|------|-------------------|------------------------------|
| Subjects x noise x scope sensitivity x frequency | 269.7534 | 18 | 14.9863 | ^b _{1.95} |
| Subjects x noise x frequency x amplitude | 394.9125 | 27 | 14.6264 | ^a _{1.91} |
| Subjects x noise x frequency x stick sensitivity | 468.2110 | 18 | 26.0117 | ^a _{3.39} |
| Subjects x noise x amplitude x stick sensitivity | 500.1103 | 18 | 27.7839 | ^a _{3.62} |
| Subjects x scope sensitivity x frequency x stick sensitivity | 445.7121 | 36 | 12.3809 | ^b _{1.61} |
| Noise x scope sensitivity x frequency x amplitude | 484.5497 | 18 | 26.9194 | ^a _{3.51} |
| Noise x scope sensitivity x frequency x stick sensitivity | 199.5423 | 12 | 16.6285 | ^b _{2.17} |
| Noise x scope sensitivity x amplitude x stick sensitivity | 230.6763 | 12 | 19.2230 | ^a _{2.50} |
| Noise x frequency x amplitude x stick sensitivity | 1336.0243 | 18 | 74.2236 | ^a _{9.67} |
| Scope sensitivity x frequency x amplitude x stick sensitivity | 682.5444 | 36 | 18.9596 | ^a _{2.47} |
| Subjects x noise x scope sensitivity x frequency x stick sensitivity | 494.6018 | 36 | 13.7389 | ^b _{1.79} |
| Subjects x noise x frequency x amplitude x stick sensitivity | 1491.7914 | 54 | 27.6258 | ^a _{3.60} |
| Noise x scope sensitivity x frequency x amplitude x stick sensitivity | 950.7665 | 36 | 26.4102 | ^a _{3.44} |
| Residual | 828.6448 | 108 | 7.6726 | |
| Total | 59053.4023 | 1151 | | |

^aP=0.01^bP=0.05

result of this relationship the gain has not only decreased but has also become more variable in the noise condition as compared to the no-noise condition.

Since tables 3 and 4 and figure 2 suggest that the basic variable of average rate of stick motion may be a pertinent variable in tracking performance, figures 3 and 4 present an analysis of the data shown in figure 2 in terms of average rate of stick motion. Figure 3 shows error as a function of rate with the presence and absence of noise and input amplitude as parameters. The figure shows that error rises as a function of rate, but independently of amplitude with noise as the only important parameter. Figure 3 also shows that when the rate reaches about 0.05 rad/sec, there is no longer a difference between noise and no-noise performance. Figure 4 shows stick movement as a function of rate and amplitude for noise

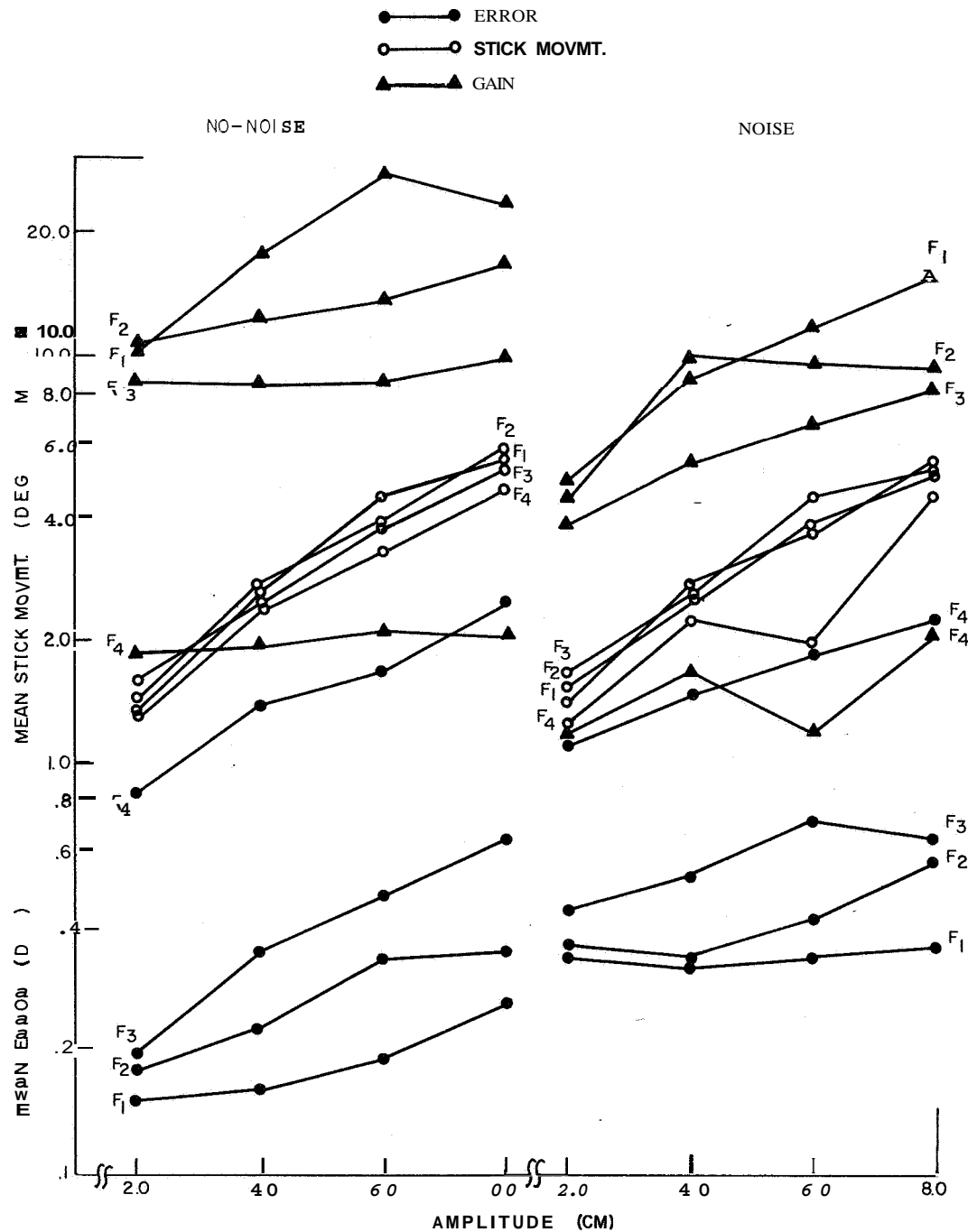


Figure 2.—Mean error, mean stick movement, and mean gain as a function of amplitude and frequency for noise and no-noise conditions.

and no-noise. Stick movement is shown to be relatively constant as a function of rate, while rising distinctly as a function of amplitude. Since rate is a combination of frequency and amplitude for a given stick sensitivity, the data points of figure 4 also indicate that frequency alone is not a factor in determining the amount of stick movement.

The foregoing analysis has demonstrated that the linearity hypothesis is untenable and that the critical tracking variable is average rate of stick motion. Further evidence for this

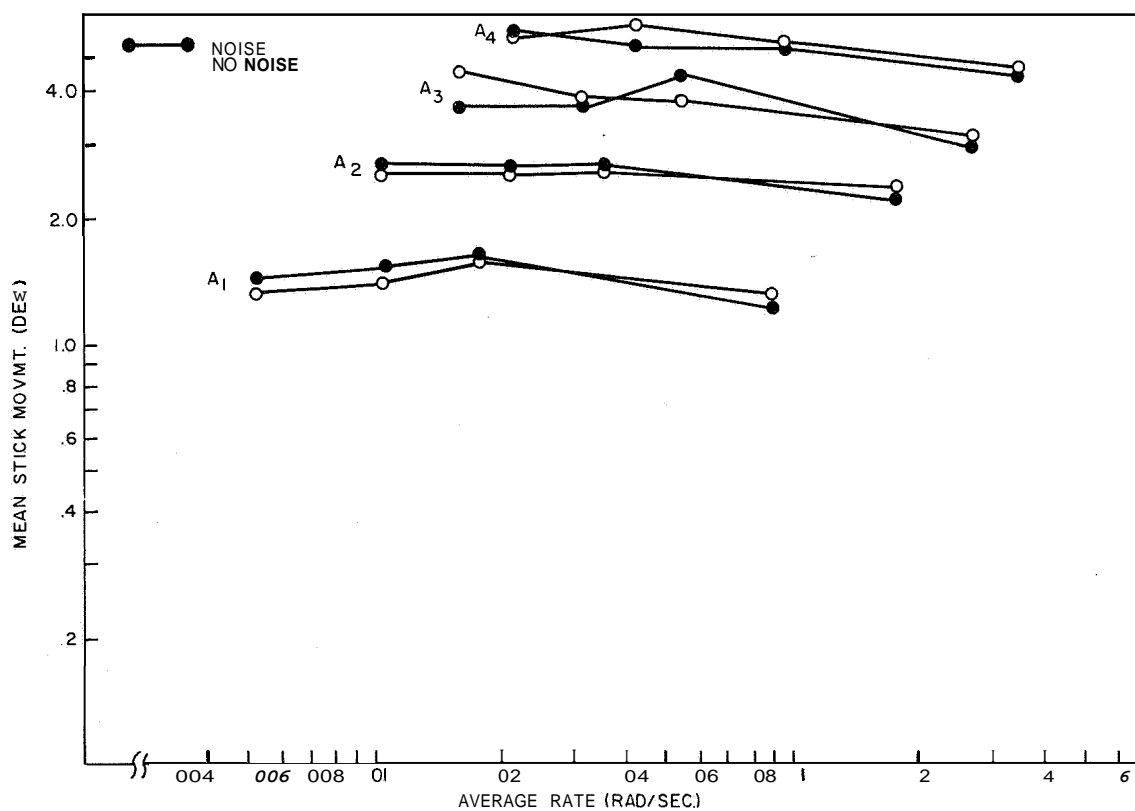
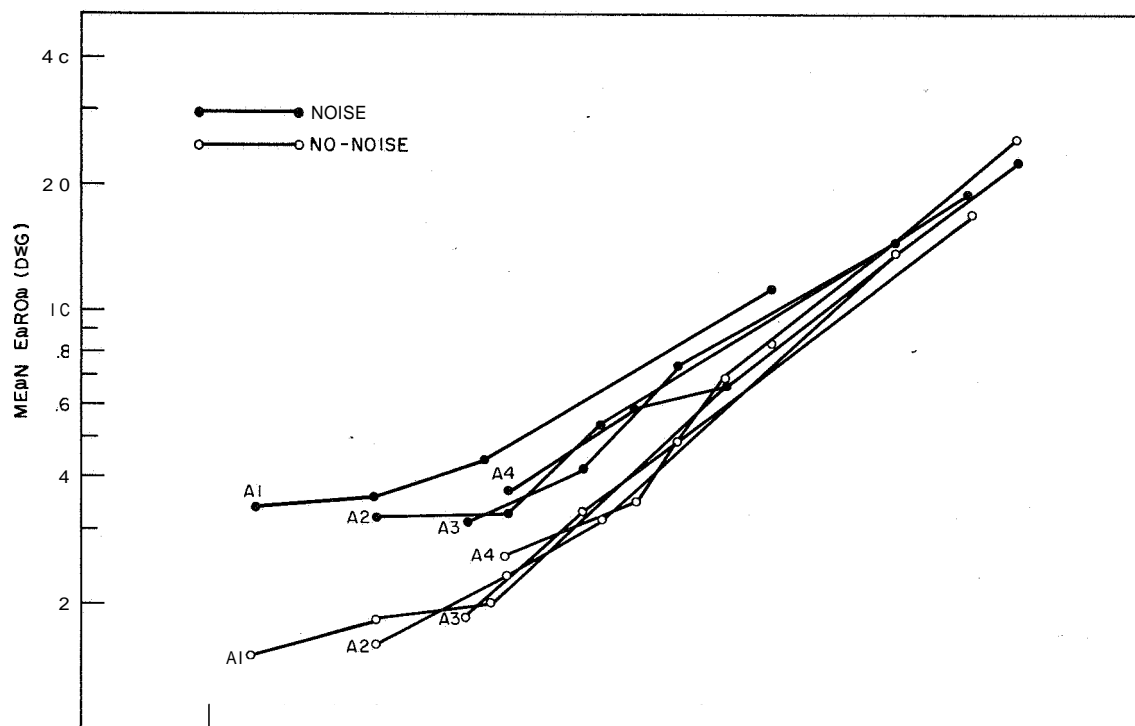
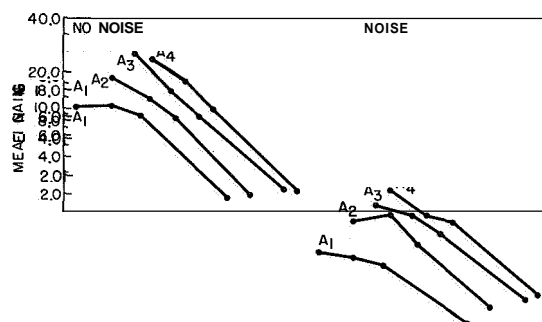


Figure 4.—Mean stick movement as a function of average rate of stick motion and amplitude for noise and no-noise conditions.



position is provided by figure 5 which shows gain as a function of average rate of stick motion, with amplitude and noise as parameters. The curves show that gain decreases predictably as a function of rate under all conditions. What mainly differentiates the noise from the no-noise con-

$$|G| = \frac{K}{\left[\sqrt{1 + \left(\frac{\omega}{\omega_0} \right)^2} \right]^m} \quad (2)$$

where

- K gain at 0 db
- ω average rate of stick motion
- ω_0 corner rate
- m measured slope

Figure 5 shows that K and ω_0 shift as a function of amplitude. The transfer function has been applied to amplitudes A₁ and A₃ to obtain theoretical curves. Figure 6 shows a fit of this function to the mean gain as a function of average rate for amplitudes 2 and 3 for the no-noise condition. Figure 7 gives the same data for the noise condition. Both figures show the empirically derived gain curves, together with the theoretical curves calculated from equation (2). The slope constant for the no-noise data is 1.0, while for the noise data it is 0.7.

In general, the fit of the theoretical curves to the empirical data is adequate and did not require the linearity assumption nor the reaction time or the intermittency hypothesis. The function is based upon the actual variables influencing behavior, that is, average rate of stick motion, noise, and amplitude of input. In these figures ω and ω_0 are the rates while K is gain at 0 db. Although this terminology is similar to the conventional servo theory terminology,

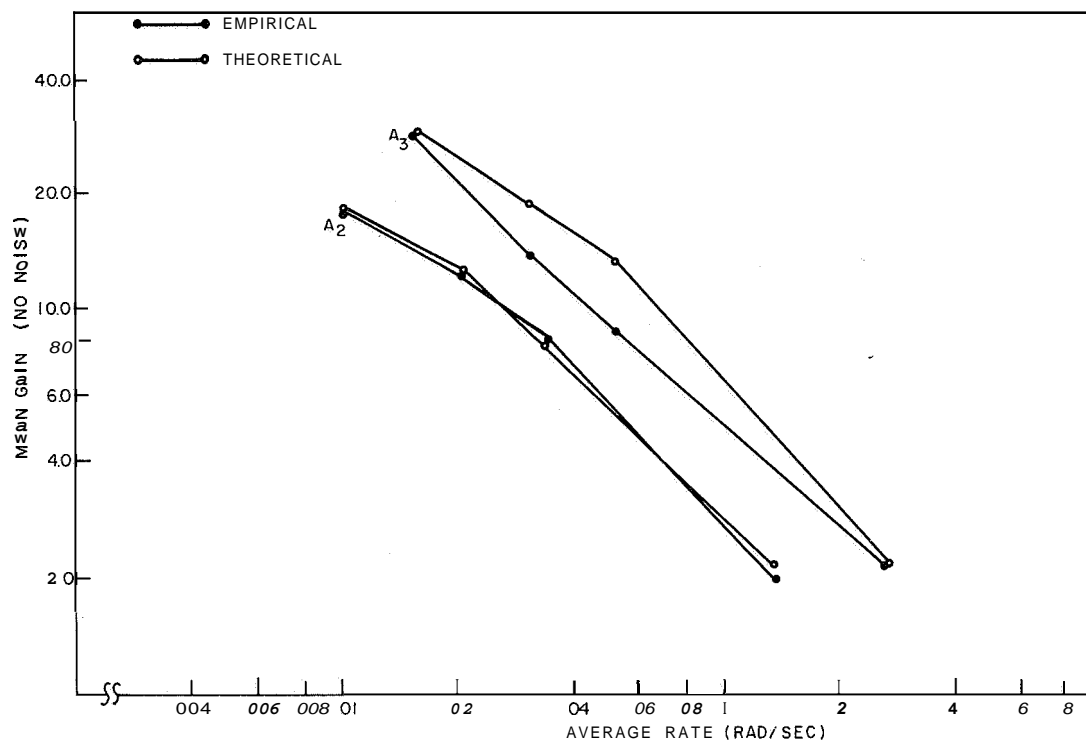


Figure 6.—Mean gain as a function of average rate of stick motion for two amplitudes, showing empirical and theoretical curves for the no-noise condition.

$$|G| = \frac{K}{\left[\sqrt{1 + \left(\frac{\omega}{\omega_0} \right)^2} \right]^{1.0}}$$

it is based primarily on the geometric relationships that exist as a result of the plot of the variables on log-log paper.

Recent work has assumed that the eye is an input device. The present study demonstrates that the eye is merely an error detector in this kind of task. Under these circumstances, the noise input merely masks the exact error that must be corrected.

The results of this study demonstrate that the performance of the human operator can be improved by reducing the rate at which he is required to move the control in those instances in which the control movement is a continuous one.

CONCLUSIONS

The hypothesis on which the research was based was confirmed by the data. From these data the following conclusions are drawn:

- (1) Human compensatory tracking performance is not linear, but varies as a function of average rate of required control stick motion. As a result,
 - (a) Tracking error increases as the average rate of required stick motion increases.
 - (b) Gain decreases as the average rate of required stick motion increases.
- (2) Imposing noise or some unpredictable portion on the signal to be tracked does not affect the shape of either the error or the gain curves, but merely serves to increase error

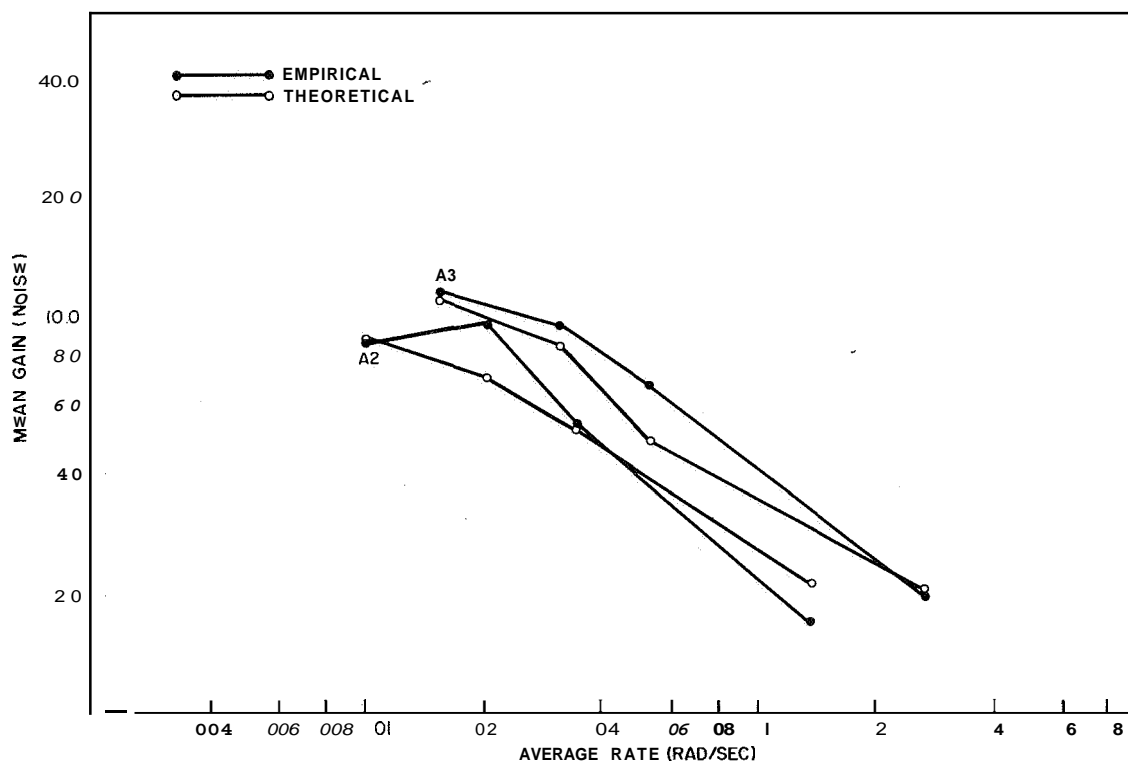


Figure 7.—Mean gain as a function of average rate of stick motion for two amplitudes, showing empirical and theoretical curves for the noise condition.

$$|G| = \frac{K}{\left[\sqrt{1 + \left(\frac{\omega}{\omega_0} \right)^3} \right]^{0.7}}$$

and decrease gain, producing the same effect as if the average required rates were in fact higher.

(3) Both the error and gain curves show that, for the particular set of variables used in this experiment the optimum range of rate at which error is low and gain is high lies between 0.004 and 0.02 rad/sec.

(4) A transfer function has been derived for operator gain, based on average rate of stick motion and on signal amplitude. The function provides an adequate fit for the empirical data.

REFERENCES

1. Beckey, G. A.: Investigation of Sampled Data Models of the Human Operator in a Control System. ASD-TDR 62-36, Feb. 1962.
2. McRuer, D.; Graham, D.; and Krendel, E.: Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15, July 1965.
3. Elkind, J. I.: Tracking Response Characteristics of the Human Operator, HFORL Memorandum, Sept. 1953.

4. Licklider, J. C. R.: Quasi-Linear Operator Models in the Study of Manual Tracking. Developments in Mathematical Psychology. R. Duncan Luce, ed., The Free Press (Glencoe, Ill.), 1960, pp. 169-279.
5. Wilde, R. W.; and Westcott, J. M.: The Characteristics of the Human Operator Engaged in a Tracking Task. Automatica, vol. 1, 1963, pp. 5-21.
6. Gibbs, C. B.: The Continuous Regulation of Skilled Response by Kinaesthetic Feedback, A. P. U. 190/53, Applied Psychology Unit, Cambridge Univ., Mar. 1953.
7. Conklin, J. E.: Effect of Control Lag on Performance in a Tracking Task, J. Exp. Psych., vol. 53, 1957, pp. 261-269.
8. Brazier, M. A. B.: The Electrical Activity of the Nervous System. Science, vol. 146, no. 3650, 1964, pp. 1423-1429.
9. Fisher, R. A.: The Design of Experiments. Fourth ed., Oliver and Boyd (London), 1947.
10. Kahn, A.; and Mazina, M.: Human Tracking Performance I: Tracking Ability as a Function of System Noise, Scope Sensitivity and Forcing Function, SR&D rep. 57-103-6-116-A-1, Westinghouse Electric Corp., 1957.
11. Truxal, J. G.: Control System Synthesis. McGraw-Hill, 1955.

9. A Comparison of Human Response Modeling in the Time and Frequency Domains

Lawrence W. Taylor, Jr.
Flight Research Center, NASA

N68

15910

Frequency and time-domain methods of analyzing human control response while performing compensatory tracking tasks are reviewed. Sample linear model results using these methods are compared and discussed. The inherent requirement of constraining the freedom of the form of the pilot models is also discussed. The constraint in the frequency domain consists of smoothing with respect to frequency; whereas, the constraint for the time domain model is more natural and meaningful in that it consists simply of limiting the memory of the pilot model. The linear models determined by both methods were almost identical.

The time domain method of analysis enables the determination of a nonlinear pilot model. The inclusion of a cubic as well as a linear term accounted for only a small additional part of the pilot's remnant and indicated that only a small portion of the total power of the pilot's output is caused by nonlinearities. The power-spectral density of an ensemble average of the pilot's output is used to determine the upper limit of the amount of power associated with a deterministic response. The indication is that more than half the remnant is stochastic when a linear model is used.

Since the NASA-University Conference on Manual Control (ref. 1) at Cambridge, Mass., February 28-March 2, 1966, additional work in human response analysis, both theoretical and experimental, has been performed at the NASA Flight Research Center. Much of this work has been part of a continuation of a NASA-USAF-Cornell Aeronautical Laboratory program to obtain pilot describing functions from both flight and simulator tests. In addition to this program, the Flight Research Center has initiated a study, under the guidance of A. V. Balakrishnan of UCLA, of nonlinear time domain methods as applied to the problem of modeling the pilot in a compensatory tracking task.

It is the purpose of this paper to assess first the frequency domain method of analysis and then the time domain analysis. A comparison of the results of the two forms of analysis applied to a linear model is made, and their advantages and disadvantages are discussed. Next, the time domain method of analysis is applied to the identification of a nonlinear pilot model. This is the first time that the nonlinear time domain method has been applied to human response data. The power-spectral density of an ensemble average of the pilot's output is used to estimate the amounts of power that correspond to linear, deterministic, and stochastic control response.

SYMBOLS

| | |
|------------------|--|
| c | pilot output (control deflection), in. |
| $\overline{c^2}$ | mean square or total power of c , in. ² |

| | |
|-----------------|---|
| E | error matrix |
| e | error, radians |
| $F[]$ | Fourier transform |
| h | time interval, sec |
| h_p | impulse response of pilot, in./rad |
| i | input (external disturbance function), rad |
| $j = \sqrt{-1}$ | |
| K | maximum value of k |
| k | index for frequency |
| L | total record length, sec |
| M | maximum value of m , $M = \frac{T_M}{\Delta\tau}$ |
| m | index for the argument of h_p |
| N | maximum value of n |
| n | index for time |
| o | linear output of pilot model (control deflection), in. |
| $R_{xy}(\tau)$ | cross-correlation function |
| r | remnant signal of pilot model (control deflection), in. |
| s | Laplace variable |
| T | one-half total record length, sec |
| T_M | maximum memory time of the pilot model, sec |
| t | time, sec |
| W_k | weighting function |
| x, y | sample signals |
| $Y_c(j\omega)$ | controlled element transfer function, rad/in. |
| $Y_p(j\omega)$ | pilot describing function, in./rad |

| | |
|----------------------|---|
| α | time variable, sec |
| P | linear correlation coefficient |
| ρ_a | average linear coherence |
| τ | argument of h_p , sec |
| ΔT | incremental value of τ , sec |
| $\Phi_{xy}(j\omega)$ | cross-spectral density of $x(t)$ and $y(t)$ |
| $\Phi_{xx}(\omega)$ | power-spectral density of $x(t)$ |
| ω | frequency, rad/sec |
| $\Delta\omega$ | incremental value of ω , rad/sec |
| $\hat{}$ | estimate |

Matrix notation:

| | |
|----------------------|------------------------------|
| $(x), \underline{x}$ | column matrix |
| $[X]$ | rectangular or square matrix |
| X^T | transpose |
| X^{-1} | inverse |
| $*$ | complex conjugate |

Numbers used as subscripts denote the pertinent term or terms of the Volterra integral series or summation.

DESCRIPTION OF EXPERIMENT

The classical experiment for obtaining data from which pilot models can be identified is illustrated in figure 1. The pilot is asked to minimize the error e displayed to him by an oscilloscope, television screen, or meter by manipulating a controller. The controller deflection c is sent to an analog computer which computes the response of the controlled element and adds to it the input disturbance function i forming an error which, in turn, is sent to the display. The signals are either processed during the experiment or recordings are made of the signals which are later processed to obtain the model of the pilot (ref. 2). Similar experiments have been performed in flight in which the pilot maneuvers the airplane (refs. 3 and 4). Most of the data analyzed in this paper were collected as part of the joint

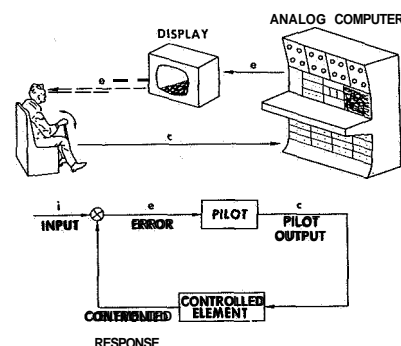


Figure 1.—Block diagram of a pilot in a compensatory task.

NASA-USAF-Cornell Aeronautical Laboratory human response studies (ref, 4) involving the T-33 variable-stability airplane and ground simulators.

DISCUSSION AND RESULTS

FREQUENCY DOMAIN METHODS.—Classically, the model of the pilot is considered to be a linear-describing function with output o plus a remnant signal r as shown in figure 2. The describing function, $Y_p(j\omega)$, can be obtained by first computing the cross-spectral density functions $\Phi_{ic}(j\omega)$ and $\Phi_{ie}(j\omega)$. The estimate of $Y_p(j\omega)$ is then given by the ratio (from ref. 5)

$$\hat{Y}_p(j\omega) = \frac{\Phi_{ic}(j\omega)}{\Phi_{ie}(j\omega)}$$

Cross-spectral density functions have generally been used instead of Fourier transforms (ref. 6) as a means of removing the bias in the estimate of $Y_p(j\omega)$ introduced by the remnant. The use of cross-spectral density functions, however, was shown in reference 7 to have no effect on the bias. The same estimate of $Y_p(j\omega)$, therefore, can be expressed as the ratio of Fourier transforms

$$\hat{Y}_p(j\omega) = \frac{F[c(t)]}{F[e(t)]}$$

and the bias in both cases is

$$\hat{Y}_p(j\omega) - Y_p(j\omega) = \frac{\Phi_{ir}(j\omega)}{\Phi_{ie}(j\omega)} = \frac{F[r(t)]}{F[e(t)]}$$

These conclusions, which were developed in reference 7, depend on the ability to express the cross-spectral density in terms of Fourier transforms. Appendix A contains a detailed mathematical proof of the expression, and appendix B presents experimental results which further substantiate these conclusions.

The frequency domain method of analysis was used in the human control response study of reference 4 in which compensatory tracking experiments were performed in flight and on simulators. Figure 3 shows some simulator results in which the values presented for $\hat{Y}_p(j\omega)$ are the means of 10 runs for pilot A and 3 runs each for pilots B and C. The vertical lines indicate the range of $\pm 1\sigma$ for each of the points. The lack of a vertical line indicates the range to be less than the height of the symbol. The input disturbance function consisted of the sum of 10 sinusoids. Values of $\hat{Y}_p(j\omega)$ were determined at the input frequencies. The use of sinusoids for the input disturbance function has the advantage of concentrating the power at several frequencies, thereby enhancing the accuracy of the estimate of the pilot describing function at the input frequencies.

If, on the other hand, a random input is used, mathematical difficulties may be encountered when using frequency domain methods. For example, if no constraint is placed on

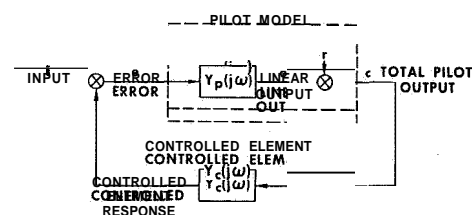


Figure 2.—Frequency domain model of a pilot.

the form of $Y_p(j\omega)$, the resulting estimate will account for the entire pilot output c erroneously indicating the remnant r to be zero. This result comes from applying the relationship

$$\hat{Y}_p(j\omega) = \frac{\bar{\Phi}_{ic}(j\omega)}{\bar{\Phi}_{ie}(j\omega)} = \frac{F[c(t)]}{F[e(t)]}$$

at all frequencies. Figure 4 shows an example of the erratic function (dashed lines) that results from such a procedure. A constraint on $\hat{Y}_p(j\omega)$ is usually provided by smoothing the values of the cross-spectral density functions or Fourier transforms, or by fairing a curve through the raw estimates of the pilot's describing function $\hat{Y}_p(j\omega)$ or both. Nevertheless, the required constraint on $\hat{Y}_p(j\omega)$ compromises one of the claims made for the analyses in the frequency domain, namely, that of unlimited freedom in model representation. Also indicated in figure 4 is the effect of smoothing on the average linear coherence ρ_a . The value is seen to change from 1.0 for the raw estimates to 0.91 for the faired case. The value of ρ_a or any other measure of the remnant has meaning only when it is connected to some particular pilot-constrained model.

If the raw cross spectral densities are smoothed, the estimate of Y_p becomes

$$\hat{Y}_p(j\omega) = \frac{\sum_{k=-K}^K W_k \bar{\Phi}_{ic}(j\omega + jk\Delta\omega)}{\sum_{k=-K}^K W_k \bar{\Phi}_{ie}(j\omega + jk\Delta\omega)}$$

This estimate can also be expressed in terms of Fourier transforms as

$$\hat{Y}_p(j\omega) = \frac{\sum_{k=-K}^K W_k F_k^*[i(t)] F_k[c(t)]}{\sum_{k=-K}^K W_k F_k^*[i(t)] F_k[e(t)]}$$

It should be noted that although the smoothing is indicated to be in the frequency domain, the same result can be obtained by averaging or truncation in the time domain.

Another point made in reference 7 is that the correlation coefficient, as it is usually defined, always has the value of unity

$$\rho_1^2(\omega) = \frac{|\bar{\Phi}_{ic}(j\omega)|^2}{\bar{\Phi}_{ii}(\omega)\bar{\Phi}_{cc}(\omega)} = \frac{F^*[i(t)]F[c(t)]F[i(t)]F^*[c(t)]}{F^*[i(t)]F[i(t)]F^*[c(t)]F[c(t)]} = 1$$

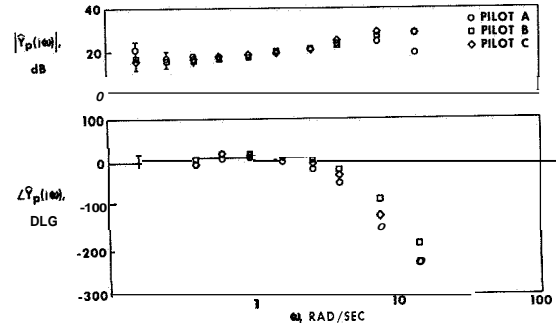


Figure 3.—Example of frequency domain model results. $Y_c \approx \frac{0.782}{s(s + \frac{1}{0.389})}$

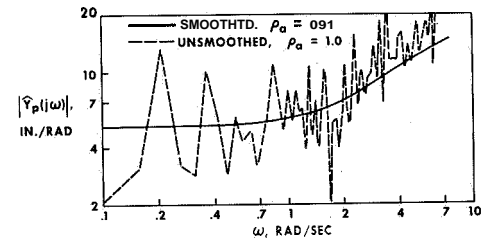


Figure 4.—Effect of smoothing the pilot describing function.

$$Y_c \approx \frac{0.782}{s(s + \frac{1}{0.389})}$$

This result is now recognized to be true in the absence of smoothing. It is now suggested that the power-spectral density be smoothed over a finite frequency bandwidth. The modified definition of ρ then takes the form

$$\begin{aligned} \rho_2^2(\omega) &= \frac{\left| \frac{1}{2K+1} \sum_{k=-K}^K \Phi_{ic}(j\omega + jk\Delta\omega) \right|^2}{\frac{1}{2K+1} \sum_{k=-K}^K \Phi_{ii}(\omega + k\Delta\omega) \frac{1}{2K+1} \sum_{k=-K}^K \Phi_{cc}(\omega + k\Delta\omega)} \\ &= \frac{\left| \sum_{k=-K}^K F_k^*[i(t)] F_k[c(t)] \right|^2}{\sum_{k=-K}^K F_k^*[i(t)] F_k[i(t)] \sum_{k=-K}^K F_k^*[c(t)] F_k[c(t)]} \\ &\neq 1 \end{aligned}$$

The linear correlation coefficient so defined is not, in general, equal to unity even for a linear system, contrary to popular belief. The coefficient does equal unity if the relationship between i and c is constant over the bandwidth, whether or not the system is linear. It is apparent, therefore, that a further improvement is needed. It is suggested that the linear correlation coefficient be expressed as

$$\rho_3^2(\omega) = \frac{\left| \sum_{n=1}^N F^*[i_n(t)] (F[c_n(t)]) \right|^2}{\sum_{n=1}^N F^*[i_n(t)] F[i(t)] \sum_{n=1}^N F^*[c_n(t)] F[c_n(t)]}$$

In this expression, an ensemble average is used instead of an average over a range in frequency. The result is that now only linear systems free of any stochastic signals will produce a value of unity, provided the input is not the same for each ensemble. In the event the input is the same, nonlinear systems will appear to be linear, with $\rho_3=1$.

LINEAR TIME DOMAIN METHOD.—Let us now consider a linear analysis in the time domain in which the output of a linear pilot model is expressed in the form (see ref. 8)

$$c(t) = \int_0^{TM} h_p(\tau) e(t - \tau) d\tau$$

Because the time histories $c(t)$ and $e(t)$ must be sampled for analysis, it is more appropriate to write

$$c(n) = \sum_{m=1}^M h_p(m) e(n-m+1)$$

or in matrix form

$$\underline{c} = E \underline{h}_p$$

where

$$E = \begin{bmatrix} e(M) & e(M-1) & \dots & e(3) & e(2) & e(1) \\ & & & e(4) & e(3) & \dots \\ & & & & e(4) & e(3) \\ & & & & & e(4) \\ & & & & & \vdots \\ & & & & & e(N-M) \\ e(N) & e(N-1) & \dots & e(N-M+1) & & \end{bmatrix}$$

$$\underline{h}_p = \begin{pmatrix} h_p(1) \\ h_p(2) \\ \vdots \\ h_p(M) \end{pmatrix}$$

$$\underline{c} = \begin{pmatrix} c(M) \\ c(M+1) \\ \vdots \\ c(N) \end{pmatrix}$$

The sampled impulse response of the pilot model, $h_p(m)$, can be obtained by using least squares

$$\underline{h}_p = [E^T E]^{-1} E^T \underline{c}$$

Inherent in the time domain representation of the pilot model is the assumption that the output at any one time is a function of only a finite time of the history of the error. The finite time period (or maximum memory) is denoted by T_M ($T_M = M\Delta\tau$) in the integral expression or by M in the summation expression for the pilot model output. For the pilot model, T_M was varied (by changing M and keeping $\Delta\tau$ constant) until it was determined that the value of $h_p(\tau)$ was essentially zero beyond about 1 second. The value of T_M selected is somewhat larger than 1 second. Figure 5 shows an example result of such an analysis. It can be seen that the model impulse response first peaks at about 0.25 second, then reverses at about 0.45 second to peak in the opposite direction at about 0.6 second, then subsides to zero. The first sample ($\tau = 0.05$ sec) is typically negative but has been faired to correspond to a pure time delay of 0.05 second. One indication of the degree to which a model represents an actual pilot is the ratio of the output of the power of the model in relation to the total power of the pilot's output. Linear pilot models will typically account for 65 to 90 percent of the total power of the output for a 4-minute run. The percentage is somewhat higher for shorter runs.

The time domain results can be transformed to the frequency domain for comparison with the frequency domain results through the use of the Fourier transform

$$\hat{Y}_p(j\omega) = F[\hat{h}_p(\tau)]$$

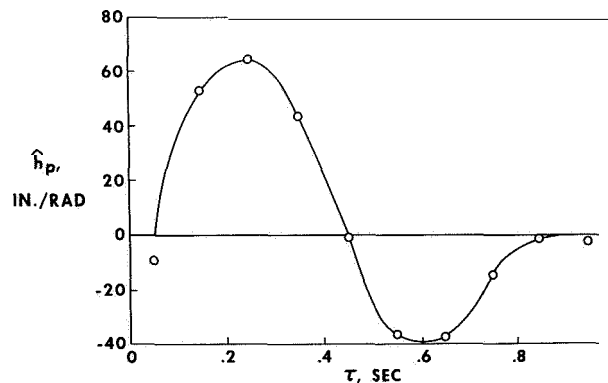


Figure 5.—Linear time domain model of the

$$\text{pilot. } Y_c \approx \frac{0.782}{s \left(s + \frac{1}{0.389} \right)}$$

It was reasoned that the pilot's control response would be symmetrical so that only the first (linear) and third (cubic) terms were used. The algorithm used to perform the analysis was again based on a least-squares solution (ref. 8).

If

$$\underline{c} = E_{1,3} \underline{h} p_{1,3}$$

where

$$\underline{c} = \begin{pmatrix} c(M) \\ c(M+1) \\ \vdots \\ c(N) \end{pmatrix}$$

$$\underline{h} p_{1,3} = \begin{pmatrix} h_1(1) \\ h_1(2) \\ \vdots \\ h_1(M) \\ h_3(1,1,1) \\ h_3(1,1,2) \\ \vdots \\ h_3(1,1,M) \\ h_3(1,2,2) \\ \vdots \\ h_3(M,M,M) \end{pmatrix}$$

$$E_{1,3} = \begin{vmatrix} e(M) & \dots & e(1) & (e(M)e(M)e(M)) & \dots & (e(M)e(M)e(1)) & (e(M)e(M-1)e(M-1)) & \dots \\ e(M+1) & \dots & e(2) & & & (e(M)e(M-1)e(1)) & \dots & (e(1)e(1)e(1)) \\ e(M+2) & & & & & & & \\ e(M+3) & & & & & & & \\ \vdots & & & & & & & \\ e(N) & \dots & e(N-M+1) & (e(N)e(N)e(N-1)) & \dots & \dots & (e(N-M+1)^2 e(N-M)) & \dots \\ & & & & & & e(N-M+1)^3 & \end{vmatrix}$$

Then

$$\hat{\underline{h}} p_{1,3} = [E_{1,3}^T E_{1,3}]^{-1} E_{1,3}^T \underline{c}$$

It is difficult to present the results of such an analysis in a meaningful form, but it is instructive to look at an example step response. Figure 7 shows the response to (1) a step of very small amplitude so that only the linear term contributes significantly to the response, and (2) a large step. The responses have been normalized to the amplitude of the step inputs

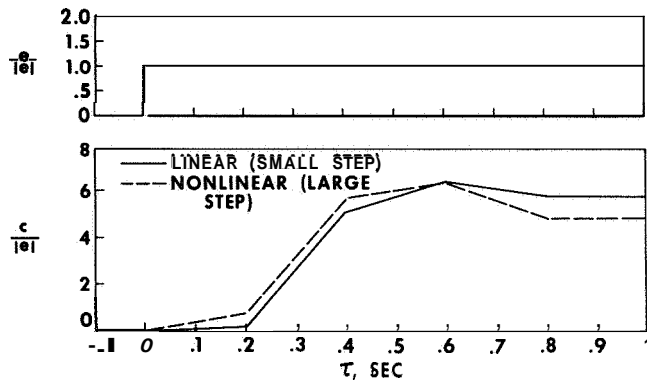


Figure 7.—Step responses of the nonlinear time domain pilot model. $Y_c = \frac{4.4}{s}$

Through the employment of the adjoint system of equations, a cubic weighting function of the form is obtained. This equation reduces the weighting function of three variables to a single

$$h_{p3}(\tau_1, \tau_2, \tau_3) = \int_0^L f_3(t - \tau) f_3(t - \tau_2) f_3(t - \tau_3) d\tau_1 d\tau_2 d\tau_3$$

function of one variable. This technique has not yet been applied to the problem of identifying pilot models, and it is not known if the reduction in dimensionality justifies the added computation required.

ANALYSIS OF THE PILOT'S OUTPUT.—It would be of interest to know what portion of the pilot's response is deterministic, but not linear, in order to assess the potential of a nonlinear pilot model in describing the pilot's output. It is known that at least part of the pilot's output is stochastic, since results of repeated experiments are never identical. It is possible to estimate the proportioning of the power of the pilot's output by examining the power-spectral density functions of both the pilot's output and its ensemble average. Both functions are shown in figure 8. The hatched peaks in the graph show the amount of power associated with a linear response at the frequencies of the input. The shaded areas show the change in the power as a result of ensemble averaging. Since the deterministic response would be unchanged by averaging, the shaded areas are an indication of the power associated with the stochastic portion of the pilot's output, which will not be accounted for by a deterministic model. The unshaded areas, then, are upper limits on the potential increase in power accounted for by using a nonlinear rather than linear pilot model.

The bar graph at the right of figure 8 shows the proportioning of the power of the pilot's output to be

to facilitate comparison. The response of the pilot model to the larger step is slightly faster, has more overshoot, and has a lower steady-state gain. The inclusion of the cubic term increased the ratio of the power of the model output to the total power of the pilot output by only a few percent. This result would indicate the remnant to be largely stochastic as opposed to nonlinear and deterministic.

If the nonlinear model is expanded to include more samples of the cubic term and higher order terms, dimensionality will become a problem. One means of reducing the total dimension is offered by Balakrishnan (ref. 8) and Hsieh (ref. 9).

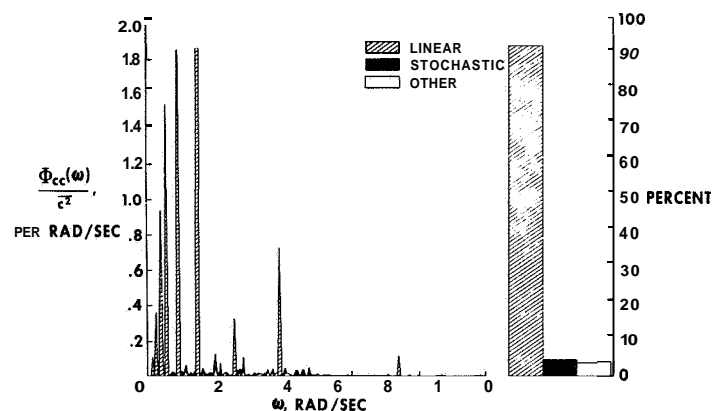


Figure 8.—Spectral analysis of pilot's output.

$$Y_c = \frac{4.4}{s}$$

91.7-percent linear (and time invariant), 4.5-percent stochastic, and 3.8-percent nonlinear and other types of responses. These results should not be generalized, since changes in the controlled element and input can cause a marked change in the proportioning of the power. It should also be noted that a small amount of power may be attributed to a nonlinearity that is significant in other aspects, such as limit cycles.

CONCLUDING REMARKS

A review of frequency and time domain methods of analysis shows that both methods require constraints on the freedom of the pilot models. The constraint in the time domain is more natural and straightforward than that of smoothing in the frequency domain. The two methods show good agreement for the linear model when the input disturbance function consists of sinusoids.

The inclusion of a cubic term in the time-domain pilot model represents the first time the analysis has been applied to human response data. For the example discussed, only a few additional percent of the power of the remnant was accounted for by the addition of the cubic term. An investigation of the power-spectral density of an ensemble average of pilot output indicates the reason to be the largely stochastic nature of the remnant. The proportioning of the power of the pilot's output appears to be about 92 percent due to linear response, 4 percent due to stochastic response, and 4 percent due to nonlinear and other types of responses.

With this step toward the application of time domain methods of analyzing human control response, the work ahead holds much promise for the determination of more meaningful and useful pilot models.

APPENDIX A — COMPARISON OF TWO ESTIMATES OF CROSS-SPECTRAL DENSITY APPLIED TO RANDOM SIGNALS

Two samples of a random signal were used to test the equivalence of two estimates of cross-spectral density. Figure 9 shows time histories of the two samples for 250 of the 400 time points used. The cross-correlation function given by

$$R_{xy}(\tau) = \frac{1}{2T} \int_{-T}^T x(t)y(t + \tau)dt \approx \frac{1}{N} \sum_{n=1}^N x_n y_{n+m}$$

where $\tau = mh$ and $h = \text{time interval}$ was computed and is plotted in figure 10. The random nature of the signals is borne out by the erratic nature of the cross-correlation function.

Figures 11(a) and 11(b) show the real and imaginary parts of the cross-spectral density as estimated by two different expressions, as follows:

$$\Phi_{xy_1}(j\omega) = F[R_{xy}(\tau)] \approx \frac{h}{N+1} \sum_{k=-N}^N e^{-j\omega kh} \sum_{n=-N/2}^{N/2} x_n y_{n+k}$$

and

$$\Phi_{xy_2}(j\omega) = \frac{F^*[x]F[y]}{2T} \approx \frac{h}{N+1} \left(\sum_{n=-N/2}^{N/2} x_n e^{j\omega nh} \right) \left(\sum_{m=-N/2}^{N/2} y_m e^{-j\omega mh} \right)$$

The values obtained by using the preceding expressions were identical except for an occasional difference in the fourth significant figure.

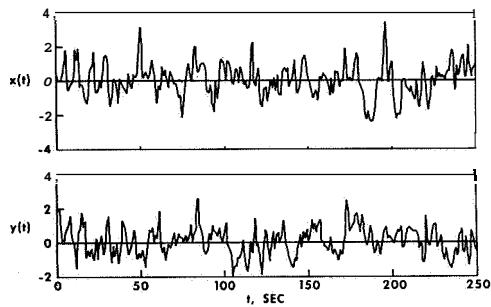


Figure 9.—Time histories of two independent random signals.

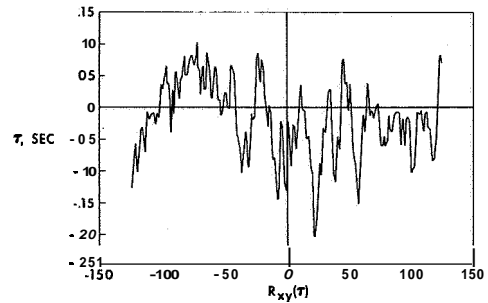


Figure 10.—Cross-correlation function of two random variables.

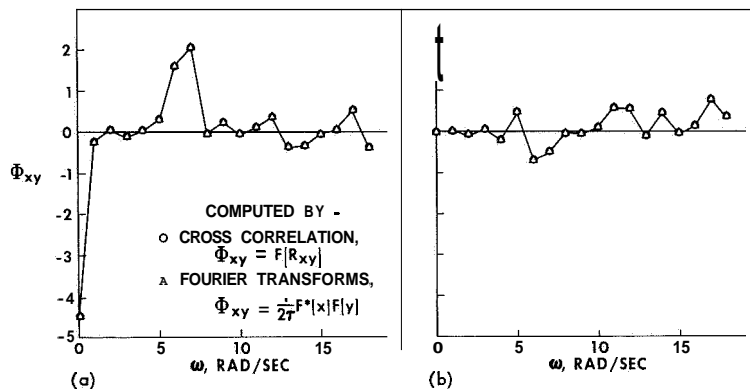


Figure 11.—Comparison of cross spectra computed two different ways. (a) Real parts. (b) Imaginary parts.

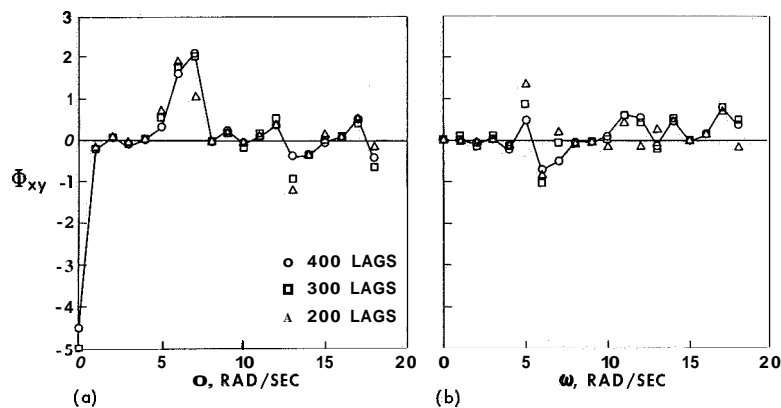


Figure 12.—Effect of number of "lags" on the estimate of cross spectra. (a) Real parts. (b) Imaginary parts.

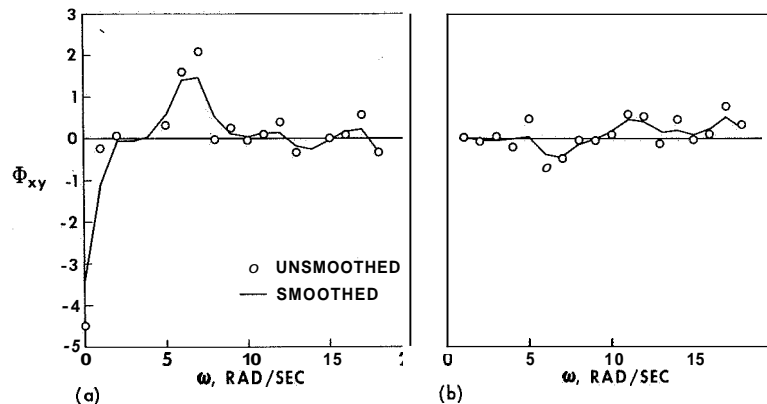


Figure 13.—Effect of smoothing on cross-spectra estimates. (a) Real parts. (b) Imaginary parts.

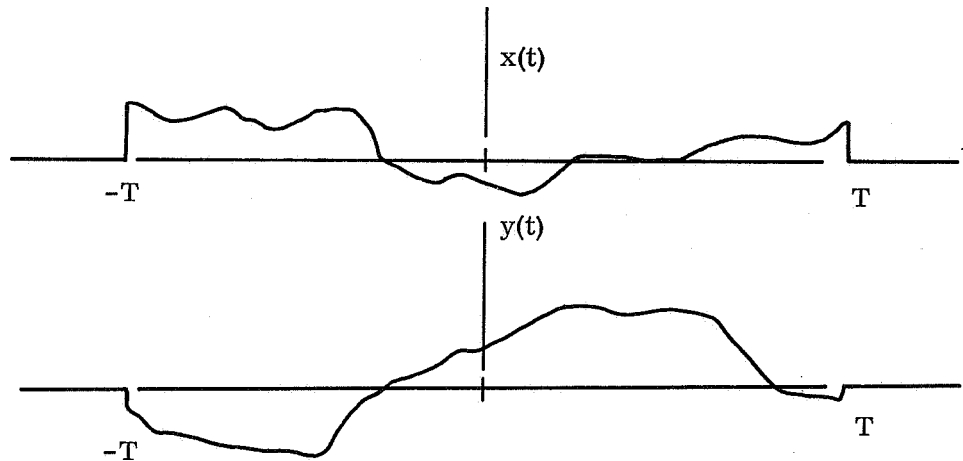
One possible explanation why calculations made by other investigators resulted in a difference for the two estimates is the effect of using a limited number of values (lags) of the cross-correlation function. Figures 12(a) and 12(b) show the effect of using fewer lags than the maximum number possible (in this case, 400). As fewer lags are used, the estimate departs from the example for 400 lags, nullifying the equivalence of the two estimates.

Another possible explanation for an apparent difference in the two estimates could be the effect of smoothing shown in figures 13(a) and 13(b). If smoothing were used for $\bar{\Phi}_{xy1}(j\omega) = F[R_{xy}(\tau)]$ and not for $\bar{\Phi}_{xy2}(j\omega) = \frac{F^*[x]F[y]}{2T}$, the difference could erroneously be interpreted as being caused by the form of the estimates of cross-spectral density.

In conclusion, the two forms of cross-spectral density estimates yield identical results if (1) the number of "lags" is not limited, and (2) identical smoothing is used in both bases.

APPENDIX B — EQUIVALENCE OF TWO ESTIMATES OF CROSS-SPECTRAL DENSITY

Consider the following sample time histories:



$$x(t) = 0 \text{ for } t > T \text{ and } t < -T$$

$$y(t) = 0 \text{ for } t > T \text{ and } t < -T$$

An estimate of the cross-correlation function can be expressed as

$$R_{xy}(\tau) = \frac{1}{2T} \int_{-T}^T x(t)y(t + \tau)dt$$

(Another estimate of $R_{xy}(\tau)$ involves division by $2T - \tau$ instead of $2T$, but the expression shown is usually more desirable.)

The Fourier transforms can be expressed as

$$F[x] = \int_{-T}^T e^{-j\omega t} x(t)dt$$

$$F[y] = \int_{-T}^T e^{-j\omega t} y(t)dt$$

The most popular estimate of the cross-spectral density function is the Fourier transform of the cross-correlation function

$$\Phi_{xy_1}(j\omega) = F[R_{xy}(\tau)] = \int_{-2T}^{2T} e^{-j\omega\tau} \frac{1}{2T} \int_{-T}^T x(t)y(t+\tau) dt d\tau$$

An alternate and equivalent estimate is the product of **two** Fourier transforms

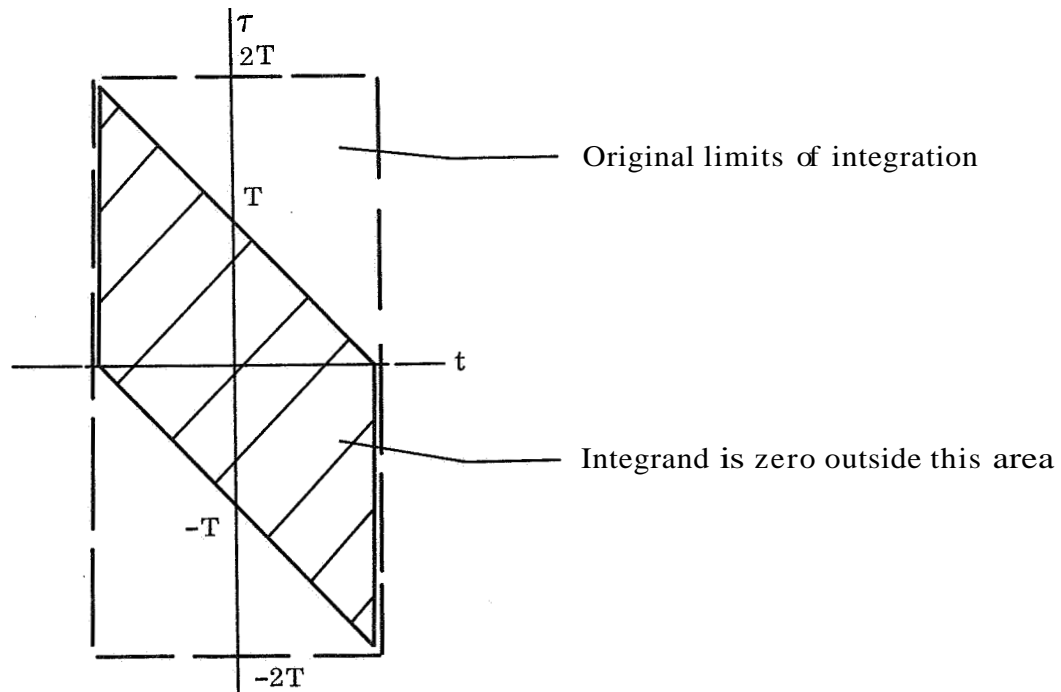
$$\Phi_{xy_2}(j\omega) = \frac{F^*[x]F[y]}{2T} = \frac{1}{2T} \int_{-T}^T x(t)e^{j\omega t} dt \int_{-T}^T y(\alpha)e^{-j\omega\alpha} d\alpha$$

in which $F^*[]$ denotes the complex conjugate of $F[]$.

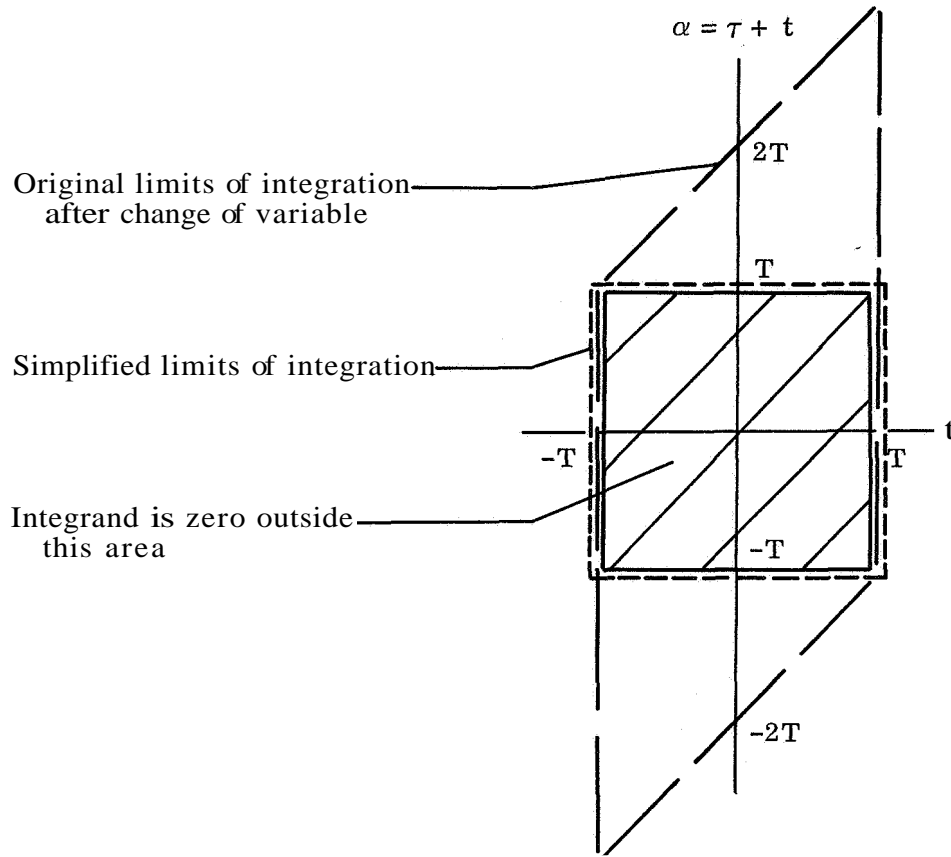
It is now desired to show that the **two** estimates are equal. This is done by substituting a change in variables and rearranging terms. Let $\alpha = \tau + t$ and substitute $\tau = \alpha - t$ in the expression of the first estimate

$$\Phi_{xy_1}(j\omega) = \frac{1}{2T} \int_{-2T+t}^{2T+t} \int_{-T}^T e^{-j\omega(\alpha-t)} x(t)y(\alpha) dt d\alpha$$

Examination of the values for which $x(t)$ and $y(\tau)$ and $y(\alpha)$ are zero enables the limits of integration to be simplified as illustrated in the following figure. Before the change of variable, the combinations of t and τ for which the integrand is not necessarily zero are represented by the shaded area. Because $x(t)$ and $y(\tau)$ are zero for $t, \tau > T$, or $< -T$, the integrand is necessarily zero outside the shaded area.



After the change in variable, the same area is transformed to that shown in the following figure:



Superimposed on the same plot are the simplified limits of integration. These limits have no effect on the integral, since both sets of limits completely cover the area for which the integrand is not zero. Therefore

$$\Phi_{xy_1}(j\omega) = \frac{1}{2T} \int_{-T}^T \int_{-T}^T e^{-j\omega(\alpha-t)} x(t)y(\alpha) dt d\alpha$$

It is now possible to rearrange the expression to get

$$\Phi_{xy_1}(j\omega) = \frac{1}{2T} \left[\int_{-T}^T e^{j\omega t} x(t) dt \right] \left[\int_{-T}^T e^{-j\omega \alpha} y(\alpha) d\alpha \right]$$

This expression is recognized to be the second (postulated) estimate. Thus

$$\Phi_{xy_1} = \frac{1}{2T} F^*[x] F[y] = \Phi_{xy_2}$$

The **two** estimates of cross- (or power) spectral density are, therefore, equal.

A similar proof of the equivalence of the two estimates of cross- and power-spectral densities is offered for the sampled time histories where

$$F^*[x] = \int_{-T}^T x(t) e^{j\omega t} dt \approx \sum_{n=-N/2}^{N/2} x_n e^{j\omega nh}$$

$$F[y] = \int_{-T}^T y(\alpha) e^{-j\omega \alpha} d\alpha \approx \sum_{m=-N/2}^{N/2} y_m e^{-j\omega mh}$$

The second estimate of the cross-spectral density is then

$$\bar{\Phi}_{xy2}(j\omega) = \frac{1}{2T} F^*[x] F[y] \approx \frac{h}{N+1} \left[\sum_{n=-N/2}^{N/2} x_n e^{j\omega nh} \right] \left[\sum_{m=-N/2}^{N/2} y_m e^{-j\omega mh} \right]$$

The first estimate may be written in the form

$$\bar{\Phi}_{xy1}(j\omega) = F[R_{xy}(\tau)] = \sum_{k=-N}^N e^{-j\omega kh} \frac{h}{N+1} \sum_{n=-N/2}^{N/2} x_n y_{n+k}$$

Then, letting $m = k + n$, and substituting $k = m - n$

$$\bar{\Phi}_{xy1}(j\omega) = \frac{h}{N+1} \sum_{m=-N}^{N+n} e^{-j\omega(m-n)h} \sum_{n=-N/2}^{N/2} x_n y_m$$

Rearranging

$$\bar{\Phi}_{xy1}(j\omega) = \frac{h}{N+1} \sum_{n=-N/2}^{N/2} e^{j\omega nh} x_n \sum_{m=-N+n}^{N+n} e^{-j\omega mh} y_m$$

Since x_n and y_m equal zero for $m, n > N/2$, and $< -N/2$, the limits can be changed to $\pm N/2$. Therefore

$$\bar{\Phi}_{xy1}(j\omega) = \frac{h}{N+1} \sum_{n=-N/2}^{N/2} e^{j\omega nh} x_n \sum_{m=-N/2}^{N/2} e^{-j\omega mh} y_m = \bar{\Phi}_{xy2}$$

REFERENCES

1. Anon.: Second Annual NASA-University Conference on Manual Control. NASA SP-128, 1967.
2. McRuer, Duane; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems. Tech. Rep. AFFDL TR-65-15, U. S. Air Force, July 1965.
3. Seckel, Edward; Hall, Ian, A. M.; McRuer, Duane T.; and Weir, David H.: Human Pilot Dynamic Response in Flight and Simulator. Tech. Rep. 57-520 (ASTIA No. AD130988), Wright Air Dev. Center, U.S. Air Force, Aug. 1958.

4. Smith, Harriet J.: Human Describing Functions Measured in Flight and on Simulators. Second NASA-University Conference on Manual Control, NASA SP-128, pp. 279-290.
5. Graham, Dunstan; and McRuer, Duane: Analysis of Nonlinear Control Systems. John Wiley & Sons, Inc., 1961.
6. James, H. M.; Nichols, N. B.; and Phillips, R. S.: Theory of Servomechanisms. McGraw-Hill Book Co., Inc., 1947.
7. Taylor, Lawrence W., Jr.: Discussion of Spectral Human Response Analysis. Second NASA-University Conference on Manual Control, NASA SP-128, 1967, pp. 403-412.
8. Balakrishnan, A. V.: Determination of Nonlinear Systems From Input-Output Data. Presented at 54th Meeting of the Princeton University Conference on Identification Problems in Communication and Control Systems, March 21-22, 1963, Princeton University, 1963, pp. 31-49.
9. Hsieh, H. C.: An On-line Identification Scheme for Multivariable Nonlinear Systems. Computing Methods in Optimization Problems, A. V. Balakrishnan and Lucien W. Neustadt, eds., Academic Press (New York), 1964, pp. 193-210.

20

2000

2000

2000

PRECEDING PAGE BLANK NOT FILMED

III. DECISION PROCESSES I

10. Information-Processing Rate as Influenced by the Degree of Response Difficulty: A Discrete Tracking Task

Daniel L. Baty
Ames Research Center, NASA

N68-15911

This study was designed to investigate the dependence of the information-processing rate on the degree of response, or task, difficulty. The degree of task difficulty was quantified for a series of discrete, random-input tracking tasks. Performance on one-dimensional (1-D) and two-dimensional (2-D) tasks was compared at equal values of task difficulty. Fourteen tasks were used: 9 with 4 target alternatives and 5 with 16 target alternatives. Six subjects performed the self-paced tasks by rapidly touching with a stylus well-defined areas as they were successively illuminated in a random sequence.

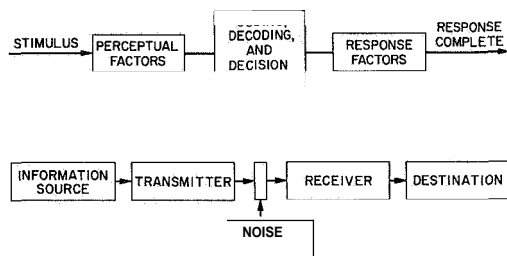
Performance in terms of the information-processing rate was primarily determined by the degree of response difficulty and the number of target alternatives. Performance in terms of the average time per response, however, was determined primarily by the degree of response difficulty. For this type of task, higher information-processing rates are possible for 2-D tasks than for 1-D tasks because, for a given number of stimulus alternatives and a constant target size, the 2-D grouping has the lower value of task difficulty.

The extrapolation of theorems from the field of information theory to the human suggests that there may be a human channel capacity which limits the information that can be processed while performing multiple concurrent subtasks. If these theorems can be applied to the human, even with modification, the systems design engineer will have a powerful new tool. In order to verify (or disprove) the applicability of such an approach, the variables that control the human information-processing rate for many different types of operator tasks must be identified and quantified.

The following three categories may help put this study in proper perspective: (a) single channel, continuous input - single continuous control, (b) multiple channel, continuous input - single discrete control, and (c) multiple discrete responses.

Continuous control to maintain acceptable vehicle position is of prime concern to mission success. Related to this first type of task is the work of Holding (ref. 1), Elkind and Sprague (ref. 2), and Crossman (ref. 3), who measured human information-processing capacity during simple continuous tracking tasks. It was pointed out by Wempe and Baty (ref. 4) that even for this simple task it was not clear which of the possible information measures is "correct." It may be that the different methods of measurement may prove useful for different purposes. Even so, each method must be related and the overall decrement from the total channel capacity due to involvement with the tracking task must be determined.

In the second type of task the operator repeatedly scans or monitors several instruments that display continuous information and takes corrective action as required. Senders (ref. 5) used information measures as a basis for modeling operator visual-sampling performance for this task. There was no selection of response in these experiments, that is,



compared. Pertinent relationships between the two studies will be discussed after the results of this experiment are presented.

The method of measuring the degree of task difficulty for this study was adapted from a study by Fitts and Peterson (ref. 9), who defined an index of task difficulty ID for a single reaction motor task as $ID = \log_2(2A/W)$, where A was the distance of hand travel and W was the width of the target to be hit. Since their conditions were two-choice tasks, only one moving distance had to be measured for any one part of the experiment, and the information value per response I_V was always equal to 1. Since they reported average times from the onset of the stimulus light to the start of hand movement, and from the beginning of hand movement to contact with the target, the average information-processing rate I_R could be estimated for each value of ID . This estimate of I_R for each condition showed that I_R was a steadily decreasing function of ID , which suggested that an index of difficulty may also have been an important variable in a study such as Klemmer's. For this experiment, Fitts and Peterson's definition of ID was adapted for tasks with higher values of I_V by substituting average hand-travel distance \bar{A} during a sequence of responses for the single distance A to obtain \bar{ID} , that is, average ID .

METHOD

A 12.5-inch-square tracking surface was divided into a 25x25 matrix of 0.5-inch-square cells, each deep enough to hold one small neon bulb below a glass cover. All cells were covered with a sheet of frosted mylar so that each 0.5-inch-square surface could be homogeneously lighted. For each of the 14 experimental conditions, an opaque mask with either 4 or 16 square holes was placed over the mylar and under the glass, making the number and location of the stimulus choices explicit. Six of the patterns were arranged as a 1-D task and eight as a 2-D task. Each mask had a different combination of hole size and spacing corresponding to a given value of \bar{ID} , as shown in figure 2. The resulting hole sizes were 0.5 inch square for conditions 1, 2, 3, 5, 7, 10, 12, 13, and 14; 0.25 inch square for conditions 4, 8, 9, and 11; and 1.50 inches square for condition 6. The term \bar{ID} is defined as $\log_2(2\bar{A}/W)$, where W is the width of the target, and \bar{A} , the average movement distance of the stylus tip for each response during the run. For this experiment, the information value of each stimulus presentation I_V is $\log_2 N$, where N is the number of stimulus alternatives, each chosen with equal probability; I_R is the average rate of processing the presented information I_V in bits/second.

Six male engineers and scientists, 28 to 42 years of age, free from physical defects, volunteered as subjects. Their task was to touch each lighted area with a stylus as quickly as possible after the onset of the light. The test equipment was designed so that the touched light would go out immediately and another randomly chosen (tape controlled) light would come on within 7 msec. Each run continued for 105 stimulus presentations. A timer automatically started after the fifth response so that all times were recorded for 100 responses. The stylus tip contained a photocell. A correct response was registered when there was a combination of photocell response and stylus contact with the board. To prevent a "sliding" strategy while going from one lighted area to another (when not adjacent), an error was counted each time the stylus made contact outside a lighted area. Ten errors were allowed per run before the run was automatically stopped and then restarted.

The experiment was conducted in a dimly lighted room to prevent spurious photocell responses. The subject sat on a high stool so that his line of sight was approximately perpendicular to the tracking surface. No results were given him during the experiment. Each session consisted of one run on each of the 14 conditions, and lasted approximately 40 minutes. First there were three practice sessions and then nine data sessions. Nine different input tapes were used for the data sessions to prevent any chance of the subjects memorizing

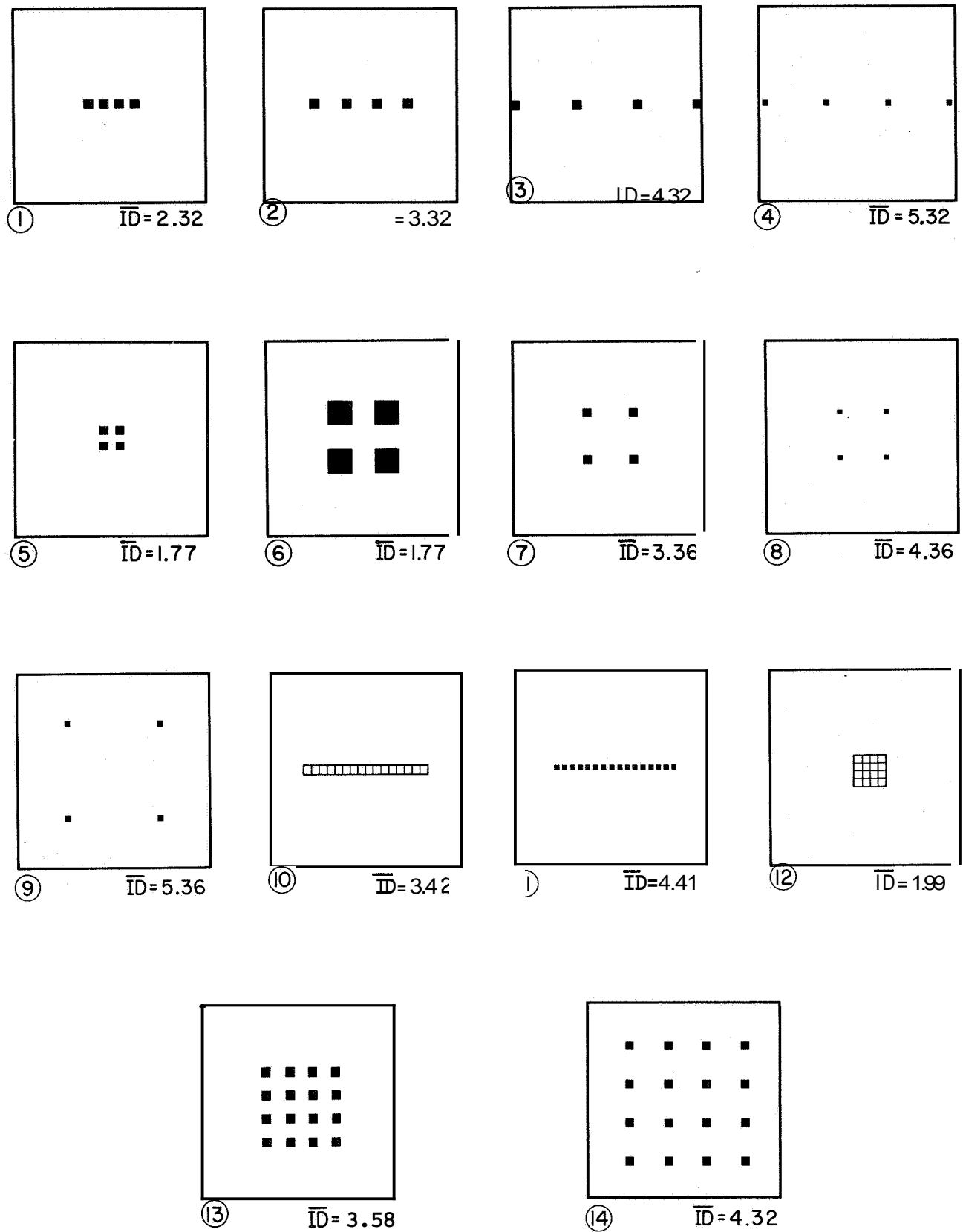


Figure 2.—Experimental conditions.

partial input sequences. The sequence of conditions within a tape was randomly chosen as was the sequence of assigning the tapes to the subjects. Within each run, the stimuli were presented randomly with equal probability with replacement; that is, a given light could repeat one or more times as the stimulus. If a light repeated immediately, contact with the board had to be broken and the light touched again. The primary instruction to the subject was to proceed as rapidly as possible after starting the run until the run automatically terminated after 105 responses.

RESULTS

The primary results of this experiment are shown in figures 3 to 10. The performance of each subject for each condition averaged over the nine experimental sessions is shown in figures 3 to 8, where average I_R is plotted as a function of \bar{ID} . By connecting the points having the same dimension (1-D or 2-D) and the same value of I_V , the important relationships for this experiment can be seen more clearly. Thus, for either value of I_V , the 1-D curve lies consistently above the 2-D curve for all subjects and for both values of I_V ; that is, I_R is consistently greater by a small amount for the 1-D task than for the 2-D task when the two are matched for \bar{ID} . There were no reversals of this result for any subject. The pattern of performance (figs. 3 to 8) is remarkably similar for all subjects even though the overall level of performance varied with the individual.

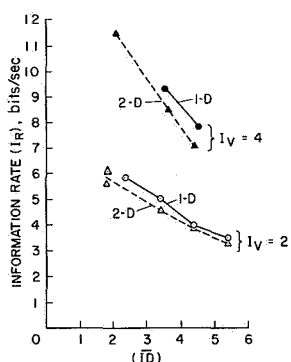


Figure 3.—The effect of \bar{ID} on I_R ; subject 1.

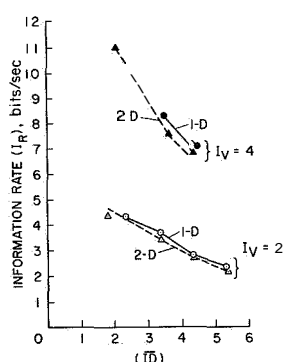


Figure 4.—The effect of \bar{ID} on I_R ; subject 2.

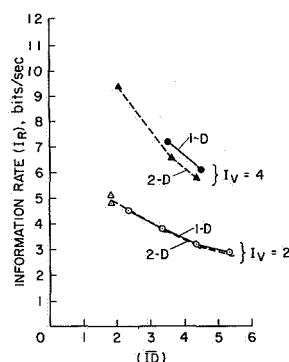


Figure 5.—The effect of \bar{ID} on I_R ; subject 3.

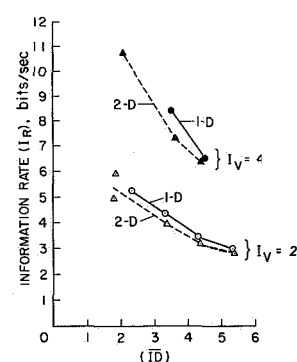


Figure 6.—The effect of \bar{ID} on I_R ; subject 4.

The average performance for all subjects is shown in figure 9. Figure 10 is a plot of the group mean times for each experimental condition (100 responses) plotted against \bar{ID} without regard for the value of I_V or for whether the task was 1-D or 2-D. Figure 10 shows that \bar{ID} is the major controlling variable in the present experiment. The average standard deviation of the run times was 3.42 percent of the total mean time per run for each condition, suggesting that the overall performance was consistent.

DISCUSSION

To illustrate further the importance of \bar{ID} in a discrete tracking task, the results of this experiment are compared with those of Klemmer's (ref. 8). Whereas only two values of I_V were used for the present experiment, with the prime concern being changes in performance

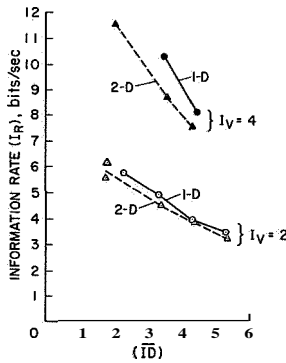


Figure 7.—The effect of \bar{E} on I_R ; subject 5.

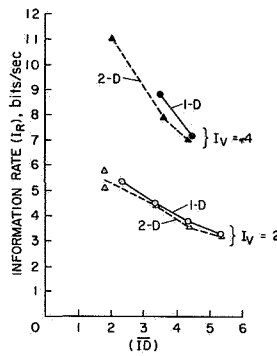


Figure 8.—The effect of \bar{ID} on I_R ; subject 6.

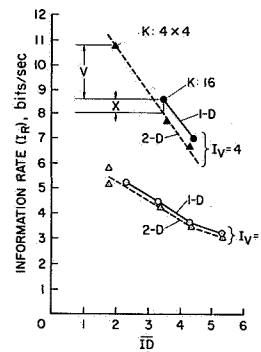


Figure 9.—The effect of \bar{E} on I_R ; mean rates from six subjects.

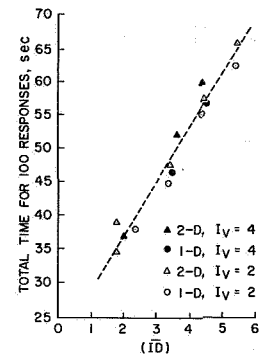


Figure 10.—The effect of \bar{E} on total response time; mean times from six subjects.

as a function of \bar{ID} , Klemmer used eight values of I_V since he was interested in performance as a function of I_V . The equipment and task configurations for the two studies were similar and two of the conditions for this experiment were designed to have exactly the same dimensions as two of Klemmer's tasks. These two conditions (configurations 10 and 12 in fig. 2) were the only two from Klemmer's experiment that had the same value of I_V for both the 1-D and the 2-D tasks. Points K:16 and K:4x4 in figure 9 show the performance for these two conditions.

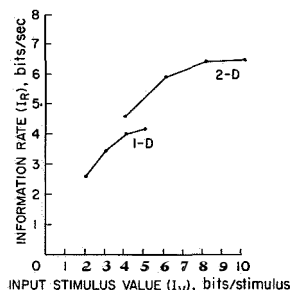


Figure 11.—Rate of information transmission for 1-D and 2-D discrete tracking tasks as a function of target position uncertainty; mean rates from eight subjects (Klemmer, ref. 8)

Klemmer's task was the same as already described for this study. His targets were all 0.5-inch squares, with no spacing between targets. Four stimulus conditions were presented as 1-D tasks ($I_V=2, 3, 4$, and 5 bits/presentation) and four conditions as 2-D tasks ($I_V=4, 6, 8$, and 10 bits/presentation). Klemmer found that I_R increased asymptotically as I_V increased toward 4.2 bits/sec for 1-D and 6.6 bits/sec for 2-D, as shown in figure 11. Klemmer concluded that these values probably represented the maximum information-processing rate for this type of task.

Since Klemmer's inputs were randomly selected with equal probabilities, and the dimensions for each task were given, it was possible to estimate \bar{ID} for each of his conditions. In figure 12, his values of I_R are plotted against \bar{ID} rather than I_V . The only points in Klemmer's experiment that had equal values of I_V , namely, points K:16 (1-D) and K:4x4 (2-D), each with $I_V=4$ bits/presentation, show that the value of I_R was greater for the 2-D task than for the 1-D task. This distance V is 0.6 bit/sec. But it is also shown in figure 12 that \bar{ID} is greater for the 1-D task than for the 2-D task. So it is not clear from Klemmer's data whether I_R is greater for his condition K:4x4 because it was a 2-D task rather than a 1-D task or because of the difference in \bar{ID} , which he did not control.

The data from the present experiment (fig. 9) show that the absolute difference in I_R between conditions K:16 and K:4x4 (distance V) is 2.2 bits/sec. A perpendicular line through K:16 intersects the fitted 2-D line of performance at the value of I_R where the interpolated estimate of performance for a 2-D task matches \bar{ID} with condition K:16. The value of I_R at this point is 0.45 bit/sec less than that

found for condition K:16, distance X. On the basis of Klemmer's results only this lower performance for 2-D tasks would not be expected. In retrospect, however, the results of the two studies are compatible, since lines drawn through points K:16 and K:4x4 in figure 12 (as shown) with slopes equal to those drawn through K:16 and K:4x4 in figure 9 result in a 1-D line above the 2-D line for Klemmer's data also. Regardless of the direction of the difference, however, this 0.45 bit/sec difference is small when compared with the absolute difference of 2.2 bits/sec between 1-D and 2-D performance for $I_V=4$ bits/presentation if \bar{ID} is not considered.

A comparison of figures 9 and 12 shows that the rates obtained for this study were much higher than those obtained by Klemmer for conditions K:16 and K:4x4. Although the equipment was similar for the two studies, there was a primary difference in the determination of when contact had been made with the board. Klemmer's subjects were required to "press down slightly," whereas for this study a slight contact closed the electronic circuit, allowing a quick touch-and-go strategy. The resulting differences in target dwell times would be directly related to a difference in I_R .

It is realized that the method used here to assign values of \bar{ID} to the experimental conditions might not reflect a "true" \bar{ID} since it was based entirely upon the geometry of the task and did not consider perceptual or physiological factors. The small elevation of the 1-D curves could well be an artifact because these factors were ignored. The general uniformity of performance, however, indicates that the \bar{ID} measure accounts for the majority of the differences between performance on the 1-D and 2-D tasks.

CONCLUDING REMARKS

The index of response difficulty (\bar{ID}), as defined in this paper, is an important independent variable to consider when operator information-processing rates for a discrete tracking task are being measured. This was illustrated by the results of this experiment and also by a discussion of the results obtained by Klemmer. It was shown that part of the results discussed in a different way by Klemmer could also be explained in terms of \bar{ID} as an independent variable.

For this type of experiment, the main advantage of the 2-D task over the 1-D task, if target sizes are held constant, is that I_V can be made much larger for the 2-D task without increasing the values of \bar{ID} . Thus, the higher information-processing rates possible for 2-D tasks are directly attributable to lower values of \bar{ID} . Information-processing rates for this kind of task are bound directly by physical response limitations. Because of inertia and energy limitations of the hand-arm response system used in this study, certain irreducible times were required for a correct response. The rest of the information-processing system had to wait while this response was being made. It is probable that the rates would have been higher if \bar{ID} had been lower.

This study points out a practical consideration when measuring human information-processing rates in a control context. The continual exchange of information between the man and his vehicle is not complete until the operator responds. Therefore, any information rate derived for the operator does consider the entire time from stimulus onset to the completion of an appropriate response. The importance of carefully considering the design of controllers and switches, and the way they are combined is thus pointed out, since they have a direct and important influence on achievable information-processing rates in pilot-vehicle tasks.

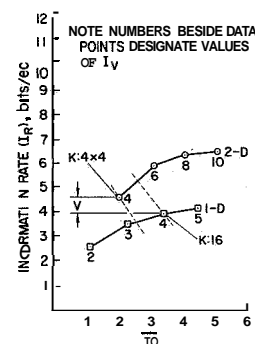


Figure 12.—Results of Klemmer's experiment (ref. 8) plotted against \bar{ID} .

REFERENCES

1. Holding, D. H.: Rates of Handling Continuous Information. FPRC-1068, Flying Personnel Res. Comm., Air Ministry, 1959.
2. Elkind, J. I.; and Sprague, L. T.: Transmission of Information in Simple Manual Control Systems. IRE Trans. on Human Factors in Electronics, vol. HFE-2, 1961, pp. 58-60.
3. Crossman, E. R. F. W.: The Information-Capacity of the Human Motor System in Pursuit Tracking. Quart. J. Exp. Psychology, vol. 12, 1960, pp. 1-16.
4. Wempe, T.; and Baty, D.: Usefulness of Transinformation as a Measure of Human Tracking Performance. Second Annual NASA-University Conference on Manual Control. NASA SP-128, 1967, pp. 111-129.
5. Senders, J. W.: The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems. IEEE Trans. on Human Factors in Electronics, vol. HFE-5, 1964, pp. 2-5.
6. Smallwood, R. D.: Internal Models and the Human Instrument Monitor. Presented at the Seventh IEEE Symposium on Human Factors in Electronics, May 1966.
7. Carbonell, J. R.: A Queuing Model of Many-Instrument Visual Sampling. Presented at the Seventh IEEE Symposium on Human Factors in Electronics, May 1966.
8. Klemmer, E. T.: Discrete Tracking in One and Two Dimensions. Rep. AFCRC-TN-56-2, Operational Applications Laboratory, Air Force Cambridge Research Center, 1956.
9. Fitts, P. M.; and Peterson, J. R.: Information Capacity of Discrete Motor Responses. J. Exp. Psychol., vol. 67, 1964, pp. 103-112.
10. Shannon, E. C.; and Weaver, W.: The Mathematical Theory of Communication. Univ. of Illinois Press (Urbana, Ill.), 1949.
11. Deininger, R. L.; and Fitts, P. M.: Stimulus-Response Compatibility, Information Theory, and Perceptual-Motor Performance. Information Theory in Psychology, H. Quastler, ed., Free Press (Glencoe, Ill.), 1956, pp. 316-349.

11. A Psychological Approach to Operator Modeling in Manual Control*

Charles R. Kelley
Dunlap and Associates, Inc.

N68-15912

The usual engineering models of the human operator employ techniques that were first developed in control engineering to describe and model control mechanisms. If there are ways in which the human operator differs from inanimate control mechanisms that are significant for representing the human operator, the usual form of engineering model would not be expected to include them. Human operators differ from control devices in that they can understand their task and its environment, can remember the past and predict and plan for the future, and can employ these abilities in control. All of these psychological aspects of the human operator are important in manual control, yet all are absent from the usual engineering models of the human operator.

A more veridical engineering model of the operator would contain the following:

- (1) A direct internal representation of the controlled element and its environment
- (2) Freedom of the internal representation from present time (fast-time operation) so that it can be used to extrapolate and predict controlled element behavior
- (3) Control action based on such predictions

Adaption in such a model would occur automatically as a consequence of changes in the internal representation, reflecting actual external changes. The internal changes would bring about changed predictions, and as a result of the changed predictions, the control action taken would be changed "adaptively."

An automatic control device containing these three features was patented in 1950 by Ziebolz. Sheridan, Johnson, and this author have employed versions of Ziebolz' controller as a human operator model. Adoptive and optimizing features of such models are examined.

The application to the human operator of the techniques of mathematical description that were developed as models of mechanical elements in a control system can be a dangerous procedure. The danger depends on how the descriptions are used, and how well their user is aware of the severe limitations of such descriptions when they are applied to man. The greatest dangers lie in attempts to apply such models to design problems in manual control. This is due in part to differences between the real-world situation in which the manual control system must function and the laboratory conditions under which data for the model were gathered. It is also due, however, to fundamental inadequacies of the mathematical tools drawn from control engineering as means of representing a human operator.

*This paper includes material from several chapters of the author's forthcoming book, *Man and the Control System*, reprinted here by permission of John Wiley & Sons, Inc.

Some of the limitations in the applicability of human operator modeling data are a necessary consequence of the simplifications imposed in laboratory work in order to reduce to a manageable number the sources of variation in data obtained. Many (although by no means all) of the investigators gathering data for human operator models have been well aware of the limitations in the application of their data to situations outside the laboratory. However, even more serious limitations are inherent in the modeling techniques themselves, rather than in their application. These are a different and more fundamental class of limitation. They are also more serious because they are not generally recognized. Three major criticisms of mathematical models of the kind that have been described for human operator representations are the following:

- (1) By their form they cannot represent the operator's highly important input data reduction function.
- (2) They fail to incorporate any explicit representation of the task performed.
- (3) The most serious criticism is that they imply that an operator's response at a point in time is a function of input at a particular point in time, failing to represent human memory, planning, and prediction processes.

Each of these is the natural consequence of the implicit assumption that descriptive techniques that were developed for automatic control mechanisms can also be applied to the human operator. Each criticism will be elaborated in turn.

THREE CRITICISMS OF TRADITIONAL HUMAN OPERATOR MODELS

INPUT NARROWNESS.—Automatic control devices are supplied with closely specified input signals and one or a few input channels. It is typical of the human operator, however, to view a rich environment, synthesizing data from many different sources received via several channels to form the basis for his response. Human operator models which by their structure require a simple input signal varying along a single dimension for each dimension of the output are more suitable to mechanisms than to men. It is this aspect of models which is associated with unrealistic impoverished displays.

The simplest sort of changes in display format can result in large changes in operator response to a given input signal content, and can easily mean the difference between stable, effective control and no control. Consider a simple example involving a three-integral $\frac{K}{s^2(1+T_s)}$ controlled element in which the operator is required to correct for step-function input changes. Employing a compensatory display, stable control with such a controlled element is extremely difficult. Early in training subjects lose control repeatedly, and even highly practiced subjects have a slow, highly inaccurate oscillatory response, with frequent large errors. Make this single change, however: continuously record and display the error signal so that the operator has a visible history of his error, and the system much more quickly becomes stable. Figure 1 illustrates typical step-function responses to the two displays. With practice, quite precise corrective responses are made to the step-function input changes with the recording display. Yet note that, considered one-dimensionally, the system and the information reaching him are precisely the same; if his input were $\theta_i(t)$ it is still $\theta_i(t)$, for no new information has been added by the changed display format. Obviously, the operator's transfer function has made a remarkable change for the better.

Transfer or describing function models are unable to represent the dramatic differences in manual tracking performance that added display dimensions and different display formats can bring about. Such models only define the relation between an input signal that varies with time along a single dimension and an output signal that does the same. But most manual control systems have more dimensions of variation in the operator's input than in his output, for

example, dimensions of position, shape, motion, and color. These dimensions allow for a multiplicity of display formats having very different effects on system performance.

INTERNAL TASK REPRESENTATION.—Human operators typically know a great deal about the nature of the task they are performing, and this knowledge reflects in their performance. The vehicle driver knows his vehicle, and somehow the characteristics of the vehicle are represented within him and condition the way he operates. In the same way, the operator knows something of the environmental factors which affect his task. The river pilot knows the river currents, the driver knows the sharp curves in a familiar road. The kinds of mathematical models of the operator that are usually employed, however, incorporate no explicit representation of the task whatsoever. True, the parameters of the model may bear a well defined relation to the task, but this is not the same as an explicit representation of it. This omission is most important in understanding operator adaptation to changes in a task. Adaptive changes assume a highly arbitrary character in models in which this adaptation is no more than the adjustment of describing function parameters. It is appropriate to say that adaptive automatic controllers increase lead, crank down their gain, etc., but such expressions represent poorly the time-varying behavior of a human operator modifying control on the basis of changes he has perceived in his vehicle or its environment.

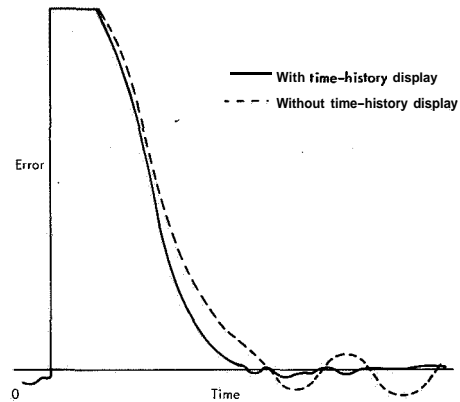


Figure 1.—Correction of a step-function error with and without a time-history display. The same controlled element, display gain, and skilled subject were employed for both curves.

POINT-IN-TIME LIMITATION.—Control mechanisms typically have no memory, nor are they able to predict. Mathematical descriptions which relate the output of such mechanisms to their inputs as of a single instant of time thus are veridical. To assume the same limitation, that is, to apply the same form of descriptive tool to man, is a serious error. After all, one of the most obvious and the one most important characteristic of man as a controller is his relative freedom from present time, due to his ability to remember the past and anticipate the future, and to employ this capacity in control.

The operator is able to look ahead, to predict, to plan, and models of the operator which leave this out grossly misrepresent him. While it is not possible to represent conscious thought per se in a model, it is possible to incorporate limited forms of "looking ahead" ("preview") and prediction in an engineering model, as will be shown. The describing function model is incapable of even representing the form of planning involved in the driver seeing the road ahead or the boat captain watching the river he is steering along, as Sheridan (ref. 1) has pointed out. Much less is the describing function model capable of representing memory or prediction of the output.

In this connection it should be emphasized that time integrals, lags, and delays cannot be considered memory for they do not fully represent the past course of events. A function cannot be reconstituted from the value of its integral at a point in time. Similarly, a derivative or lead term is not equivalent to a prediction for the same reason that a rate indicator is not a predictor display and is not equivalent to the preview of the input.

The foregoing are three of the more crucial of the inadequacies of the describing function and certain other models of the human operator. There are other corollary points as well. Table 1 summarizes the principal points made.

TABLE 1.—COMPARISON OF DESCRIBING FUNCTION MODEL
WITH ACTUAL HUMAN OPERATOR

| Describing function model | Human operator |
|--|---|
| Input narrowness | |
| (1) Input has same number of dimensions as output
(2) One display
(3) Assumes impoverished display format (compensatory or pursuit tracking) | (1) Input typically has more dimensions than output; operator reduces data
(2) Multiple displays
(3) May use highly sophisticated multidimensional displays (contact analog, predictor display, or direct view of environment) |
| | |
| (1) Restricted to present instant of time
(2) Cannot remember; can only summarize signals via integration (lag)
(3) Cannot predict input or output; response is an arbitrary weighting of error, lead, and lag terms | (1) Response based on remembered past and predicted future
(2) Can remember, modify response or change internal task representation in consequence of past experience
(3) Can predict; response can be formed and modified to minimize future (predicted) error. Can preview or anticipate input as well as predict output, and plan response based on both of these "excursions from present time" |
| | |
| (1) Cannot predict; instead responds only to derivative(s) of output (lead)
(2) Cannot adapt to changes in task save through arbitrary parameter adjustments | (1) Can extrapolate present conditions into future through internal model of task, going beyond derivatives or lead terms in response
(2) Veridical changes in internal representation of task result in changed predictions, and hence, a different nonarbitrary form of adaptation |

PREDICTIVE AND PREVIEW MODELS

Most engineering models of the human operator include neither an explicit representation of the task (i.e., of the controlled element and the input to it from the environment) nor an explicit prediction of system output, although both of these are characteristic of human operation. However, a form of model which did was proposed more than fifteen years ago,

not by an engineer studying manual control, but by one interested in developing more effective techniques of automatic control. His device can be referred to as the "Ziebolz controller."

The term "Ziebolz controller" is used here as a generic term to describe all automatic control systems built around a fast-time predictive analog or model of the system under control.¹ As such, it is not limited to the particular forms described in the original Ziebolz patent² and the subsequent Ziebolz and Paynter article (ref. 2).

Ziebolz and Paynter contrasted automatic and manual control, stating:

...If we compare this somewhat restricted control operation, [of an automatic controller] however, with that applied by an intelligent and untiring operator, we note a basic difference. The operator tends to make his observation intermittently, then—through a conscious or subconscious process of computation—he decides on direction, magnitude and method of corrective action. Observing the results, he repeats his computation process and often discounts his previous (perhaps erroneous) action.

...However, a significant point should be observed—the method of control just outlined is not based on present error but on future error and the process of computation is not a continuous one but a repetitive one, starting with a clean slate in intervals of time,...

The authors go on to present the concept of two time-scale predictive controllers that institute corrections as a consequence of the predicted future, not merely of the present and past. In this respect, these controllers resemble the human operator more closely than any of the usual forms of human operator models (e.g., describing function models).

SHERIDAN'S PREVIEW MODELS

One investigator in the field of human operator models who has been keenly aware of the inadequacies of the describing function approach, especially its inability to represent human ability to predict and to plan ahead, is T. B. Sheridan of M. I. T. Sheridan and his colleagues have, for several years, been concerned with problems of "looking ahead" in manual control, of precognition and input preview, and with developing human operator models that incorporate the ability to look ahead (ref. 1). His work has gradually expanded to incorporate, in addition to input preview, a concern with the prediction of output, including the use of a simple Ziebolz controller type of circuit to develop the output prediction (refs. 3 and 4). The Sheridan models are thus in accord with the view of the human operator that has been developed in these pages. They have the feature, unusual for such models, of attempting to represent the crucially important prediction and planning aspects of manual control.

Sheridan demonstrates the inadequacy of "instantaneous error tracking" models of the operator to represent control situations involving preview. The "instantaneous error" model cannot respond to a step-function input change until after the step-function arrives, for example, while a preview model can anticipate and begin responding in advance. Sheridan adds to his concern with preview the constraint that error is unequally important at different times in most tasks, and a good model will weight error according to its importance. Thus

¹The Ziebolz controller includes a "fast-time model" of the system under control, while the entire Ziebolz controller can be considered as a "model" of the human operator. To avoid confusing the two uses of the term "model," the fast-time portion of the Ziebolz controller will sometimes be referred to as a fast-time "analog" of the system under control.

²Ziebolz, H. U. S. Patent 2,712,414, filed November 18, 1950.

Sheridan's models incorporate preview, error-weighting by "importance function" and, in one case, output prediction.

OTHER PREDICTIVE MODELS

The Sheridan models of the operator have focussed on the problem of preview and looking ahead, employing the Ziebolz controller approach as an adjunct to one of three preview models. The predictive approach based on the Ziebolz model and the preview and planning approach of Sheridan's are fully compatible, representing, as they do, attention to future output and future input, respectively. This author's work on models, conducted concurrently with Sheridan's studies, has focussed primarily on the Ziebolz controller itself as an operator model, with only slight attention to the problem of preview (ref. 5). It has, however, dealt with some more complex applications of the fast-time predictive approach, including both the original double-loop Ziebolz controller, and a newer single-loop version.

DOUBLE-LOOP ZIEBOLZ CONTROLLER.—Figure 2 is a double-loop version of a Ziebolz controller that forms an interesting human operator model. The controller is much as described in the original Ziebolz patent and in the Ziebolz and Paynter article (ref. 2). The compensation here represents a learned and automatic response set of the operator, which is adjusted by an "evaluator" on the basis of fast-time analog-generated predictions. This model can thus be satisfied by an "automatic" controller as an inner loop, the parameters of which are adjustable on the basis of predictions generated repetitively by a fast-time analog in an outer loop. This arrangement resembles, in a general way, that of many adaptive control systems in which an outer adaptive loop changes the parameters in an inner loop. However, the outer loop is here referred to as "**optimizing**" rather than "**adaptive**." The term "**adaptive**" is reserved for the

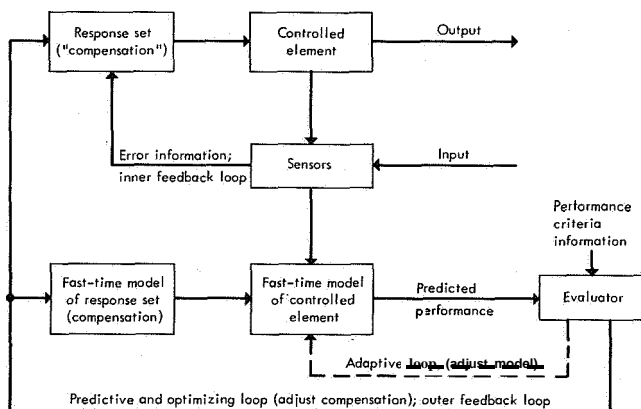


Figure 2.—Double-loop Ziebolz type controller as a human operator model. Prediction affects outer loop only.

dotted line, which varies, not the compensation parameters, but the characteristics of the fast-time analog. This human operator model is of a mixed type in terms of intermittency, since the inner loop would function continuously, the outer periodically.

The model becomes more specific when the characteristics of the various blocks are defined. The inner loop characteristics are critical. One approach is to assume that the operator's response set, or "compensation" takes a form such as that described by the familiar Laplace equation

$$H(s) = \frac{K e^{-0.2s} (1 + T_L s)}{(1 + 0.1s)(1 + T_I s)}$$

where $H(s)$ is the transfer function of the operator's response set in the inner loop, and in which K , T_L , and T_I are adjustable gain, lead, and lag coefficients, respectively.

In such a formulation, much of the control function is carried on in the inner loop, and the outer loop need only produce occasional adaptive or optimizing adjustments in the compensation terms based on predicted criterion functions. However, much more of the control function can be absorbed into the outer loop if a simpler inner loop expression is assumed. For example, if the inner loop is assumed to contain only delay and gain terms, then stability as well as adaptation becomes dependent on the outer predictive loop. The predictive process rather than the response set then provides the operator's "compensation" for controlled element and forcing function dynamics. This reduction of the function performed by the inner loop causes the model to resemble the single-loop Ziebolz controller in which the focus is entirely on the predictive process, the inner loop having been eliminated entirely.

A SINGLE-LOOP FORM OF ZIEBOLZ CONTROLLER MODEL OF THE OPERATOR.—

There is no evidence that a process of control system "compensation" described by an equation such as

$$H(s) = \frac{K e^{-\tau s} (1 + T_L s)}{(1 + T_N s)(1 + T_I s)}$$

actually goes on within the human operator, even though the operator's performance can be matched closely in many tracking systems by a device based on such an equation. Time delay τ and neuromuscular lag T_N do seem to represent real boundaries of human performance in tracking a random signal, and thus a rationale is provided for including something corresponding to them in a human operator model. However, a process different from gain, lead, and lag adjustments might better account for adaptive human operator performance.

Prediction in the model of figure 2 had an "evaluative" and an "adaptive" function, but was not a direct part of the inner control loop. It appears typical of the human operator in many control tasks that his immediate (inner loop) response is predicated on its predicted effect. This is only to say, for instance, that the driver turns the wheel so that his car will miss an obstacle ahead. A variation of the Ziebolz controller in which predictions are formed and employed in the inner control loop appears to be needed.

One such system has been reported in another context (ref. 6). In it, fast-time predictions formed by an automatic predictive controller were made repetitively at a high rate, based on the assumption that existing control action was reduced to zero. These "zero-correction" predictions were sampled, and the sample values were used directly for control. The predicted "zero-correction" error at a point in the future was employed directly as a control signal by scaling it and inverting its sign. The control signal was thus kept inversely proportional to the predicted error that would occur T seconds in the future were no further corrective action taken. By appropriate choice of prediction span T and of gain, stable diving control of a simulated submarine was achieved. Figure 3 is a diagram of such a system. Because future error is sampled in such a system, it may be desirable to smooth it by means of a filter.

The diagram of figure 3 shows how it is possible to incorporate the fast-time analog and its predictions into the inner control loop. Instead of two processes, a control process (inner loop) and a prediction process (outer loop) to adjust the control process, the prediction process itself is used for control. The control signal is thus synthesized from predictions. Because prediction is intermittent, frequency response might be reduced in comparison with the previous double-loop structure, since the control signal is formed through the sampling circuit. This incorporation of the predictive circuit in the inner loop may, nonetheless, provide a more veridical representation of the operator, including his poor frequency response

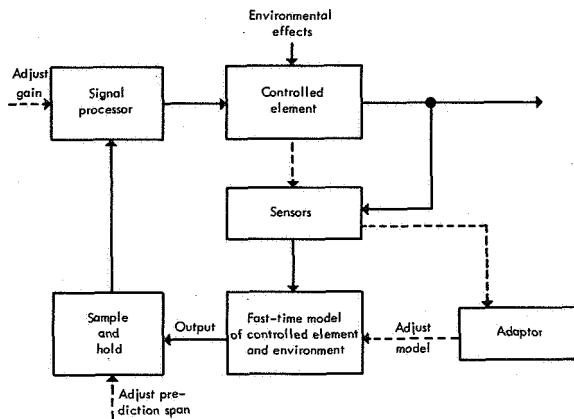


Figure 3.—Ziebolz type controller in which predictive control is exercised in inner loop.

characteristics. The rationale of the device is that it represents an operator who predicts "unless I do something, I'm going to be this much in error in T seconds,"—and who introduces a correction proportional to this predicted error.

The operator model of figure 3 is (like any device employed as a human operator representation) an oversimplification. A major weakness is basing the response on predicted error at a single future point. The operator seems more typically to develop and refine iteratively a longer segment of his response derived from a corresponding longer segment of a predicted output trajectory. A human operator model that might better represent the typical way of forming control movements would (1) anticipate future error, (2) form a response designed to reduce it, (3) try the re-

sponse out via the fast-time model, (4) assess the future error trajectory that results, and (5) form a corrected response, which might again be tried and assessed (steps 3 to 5), repeating as needed until the response is executed.

To mechanize a model which operates on this iterative basis, a control response programmer is needed to store trial programs in the process of development. Experience doubtless provides a learned repertory of responses which may at least be the starting points for many programmed responses. These responses, initially approximate, would become successively refined as prediction and plan move from future to present. The rationale for such an iterative model would be the operator who plans a control action, predicts "if I do this, then my predicted path will be—," who then corrects the planned control action as a function of the deviation of predicted and desired path, and then tests the corrected control action, etc.

Ziebolz controller technology is not yet advanced to the point of providing a working technique for refining controller response via successive iteration in a way resembling that of the human operator, although such a development is well within the state of the art in engineering. Perhaps it will not be long before such a development is made.

ADAPTIVE HUMAN OPERATOR MODELS

Most human operator models have been developed to represent human control response in stationary tasks. In recent years with the growth of interest in adaptive automatic and manual control systems, there has been a mushrooming of interest in adaptive human operator models.

The appropriate place to begin a comparison of adaptive human operator models is by a careful description in observational and psychological terms of the process of adaptation by a human operator. Then it will be possible to compare the process of human operator adaptation with the adaptive process as it occurs in various human operator models.

HOW THE HUMAN OPERATOR ADAPTS.—The basic predictive theory of manual control has been developed and presented elsewhere (ref. 7). It can be summarized as follows: the human operator exercises control by foreseeing alternative possible future states and the actions on his part that are required to bring them about, and selecting the future state of his

choice, invoking whatever criterion **or** criteria for choosing that he may deem appropriate; he is able to foresee possible future states by virtue of a dynamic internal model of the situation, created in consciousness, by means of which he can anticipate what will happen as a consequence **of** various actions he can take.

Learning on the part of the human operator then consists largely of building and refining the internal model by means of which prediction is made. Adaptation is the appropriate modification of the model in the face of objective changes in the control situation. This may include changes in the controlled element **or** in the environment. To adapt, the operator must become aware of a change in the situation which calls for an adaptive response. There are **two** categories of information that he can employ:

(1) He may receive direct information about such changes, information gathered either via his unaided senses **or** through sensing instruments.

(2) He may infer such changes by observing departures from expectancy in the system output.

He may employ both types of information, of course, and the combination can be thought of as a third type of adaptation. Whether the information bringing about adaptive changes is "direct" **or** "inferred," or a combination, the end result of the adaptive process is the same, that is, an appropriate change in the operator's internal model of the control situation. The adaptive process, however, may be quite different in the three cases, as will be seen.

Implicit in all adaptive changes is a criterion **or** criteria of performance to be met. However, when the criterion of performance rather than the controlled element **or** the environment changes, the resulting modifications in control can be considered to be due to "optimization" rather than "**adaptation**," and will be discussed in a later section.

Adaptation via "Direct" Information.—The operator's internal model represents the way in which the controlled element responds to actions he takes. Changes in the controlled element **or** its environment that affect control may be registered on sensing elements and conveyed to the operator. The sensing elements may be the sense organs of the operator **or** another observer, **or** they may be mechanical sensing devices. The only requirement is that the appropriate information be conveyed to the operator to enable him to make the changes called for in his internal model. When the sensing elements are external to the operator, the display **or** other communication of information to the operator is an important aspect of the adaptive process.

Consider a few examples of adaptation via "direct" information. A driver sees that the road ahead is icy; a sailor observes that a squall is coming up; a transport pilot is told how much added load he has taken on at a stop. Each operator incorporates the information into the internal model on which his control behavior is based. Predictions from the model then reflect these changes, and control is modified appropriately. Thus, learning that his airplane is heavily loaded revises the pilot's internal model of the control response of his aircraft; his model-based prediction of the airspeed at which lift will equal weight is changed appropriately. In this way the human operator adapts his control behavior in response to direct information.

When adaptive information is directly obtained, it may bring about adaptation before control has been affected by the change. In this respect, adaptation via direct information may be predictive, that is, the operator may learn about changes to be adapted to before they affect control. In the above examples, the driver may see that the road ahead is icy before he reaches it; the sailor may note the squall coming up and trim his sails before it hits; and the pilot may learn of the change in the weight of his aircraft before he takes off. Initial adaptive adjustments are, in such cases, made in advance. Prediction is possible because adaptation is independent of the behavior of the controlled element. Because no feedback of information from the system output is required, adaptation via direct information is "open loop" and may (often) be preprogrammed.

Adaptation via Inferred Changes.—The second major class of adaptive changes in manual control is one in which the operator has no direct information about changes in the control situation, but must learn of them through changes in the response of the controlled element. In this case, adaptation can never occur ahead of the change adapted to, as the need for change cannot be known in advance, and cannot be preprogrammed. These inferred changes are closed-loop adjustments. They are generally made on a trial basis and so are revisable.

The key source of information in adapting to inferred changes is a departure of the behavior of the controlled element from expectancy. The operator expects (predicts) a certain desired response in consequence of his internal model of the controlled element and its environment. When the response that obtains is different from that which was predicted, it indicates that the model is in error, and that adaptive change in the model is necessary. The most important single item of information for human operator adaptation is the difference between what is predicted to occur as a consequence of the operator's control actions and what actually does occur.

To illustrate, the difference between the response a driver expects and what he experiences may tell him that a tire is soft, a road surface loose, or that an unfamiliar vehicle oversteers, has brakes that fade, etc. Adaptation is likely to be tentative at first, since there may be other possible explanations for the vehicle's response. Additional feedback may confirm or deny an operator's "adaptive hypothesis."

Unlike adaptation to direct information, adaptation to inferred changes must of necessity lag the changes adapted to. It is only as a result of error in prediction that the need for such adaptation can become known. The error must therefore have existed prior to adaptation. The success of adaptation is then dependent entirely on how well the error is diagnosed.

The diagnostic problem in adaptation via inferred changes may be severe. The problem for the operator is that his information is negative. He knows that his internal model failed to give him accurate predictive information, but does not know why, or what the true situation is, unless he is able to infer it correctly. He tries to revise his model in a direction that would have produced the correct prediction, and is at the same time reasonable in terms of the operator's overall knowledge of the situation. Depending upon the confidence of the operator that the model revision is correct and upon the consequences he anticipates if he is wrong, he may accept the changed model only tentatively, and he may initiate tests to see if it is correct.

Testing his internal model is often an important part of adaptation through inferred changes in the control situation. The pilot sometimes "dithers" the stick of an aircraft to see if the plane is responding in the way he expects. The driver may test the brakes of his car after going through a deep puddle for similar reasons. Have his brakes washed out? If so, he must adapt to the change, and modify his driving to conform to the changed situation, or, more accurately, to his changed internal model of the situation. It is interesting to note that test signals also play an important role in many automatic adaptive systems.

Adaptation via Combinations of Direct Information and Inferred Changes.—The operator owes his excellence as an adaptive controller to his ability to obtain and employ a wide variety of information in the adaptive process. He can observe, infer, and correlate information from diverse sources. He can form hypotheses and test them via direct observations or via test signals, or via combinations of the two.

More often than not, the need for adaptive change is inferred from the behavior of the controlled element. Once the need for change is recognized, the operator goes into an active search for information. He is likely first to check all available direct information, for example, observing the environment and controlled element directly, and via indicators

attempting to find clues that would explain the change from expectancy in the behavior of the controlled element.

If the operator discovers with some certainty what has changed, the adaptive process proceeds smoothly. The operator changes his internal model of the situation to conform to the change, and his model-based control behavior is modified accordingly. However, if the operator is uncertain about the change he is dealing with, diagnosis and testing become important. The operator may employ deliberate test signals, or he may simply observe the response of the controlled element to no input or to normal inputs. In either case, he is intent on observing departures of the controlled element from the response his model would predict. It is these departures that make possible adaptive change in the absence of other evidence of changes in the control situation.

The "programing" of adaptive changes in response to observed or inferred changes in the control situation is greatly influenced by past experience. Is the changed situation one with which the operator is familiar, and knows how to cope with effectively? If it is, then it is only necessary that the operator be able to bring his experience to bear; the primary problem is one of diagnosing the change properly. If the change is outside the operator's experience, diagnosis is not enough; the operator must also learn to cope with the change. Such a case should not be treated as simple adaptation, as it is, in reality, a matter of adaptation plus learning. Learning is likely to be a much slower, more difficult process than adapting from one already-learned mode of control response to another. However, adaptation frequently, perhaps usually, incorporates some degree of learning. If learning consists in building an appropriate internal model of the control situation, and adaptation consists in modifying the model appropriately to reflect changes in the situation, then closely related internal processes must be involved.

Criteria Governing Human Operator Adaptation.—Control is always governed implicitly or explicitly by performance criteria. It is such criteria that determine whether or not control is being exerted at all, let alone whether it is being exerted well or badly. In adaptive control, criteria take on a special significance. How is it to be determined whether or not the operator has adapted successfully? The answer to such a question requires that criteria of system performance be examined.

A primary virtue of the human operator is his ability to understand and evaluate complex criteria and to modify his control behavior appropriately as a result. He may not control "optimally" in the sense of finding the solution to a mathematically defined criterion functions manual control that is "optimal" in this sense is seldom called for. What the operator is able to do is exercise judgment with respect to the many different criteria relevant to human action. Consider the criteria governing the automobile driver with these different goals:

- (1) To get a woman in advanced stages of labor to the maternity hospital
- (2) To conserve gasoline, because he may otherwise run out before reaching a service station
- (3) To drive a visitor from out of town on a sightseeing trip
- (4) To drive to the garage with brakes that are severely defective
- (5) To test the performance of a used car he is considering for purchase

These situations embody the types of criteria human operators are able to cope with effectively. It is evident that driving performance in the above five instances would be appropriately different, even though it involved the same operator driving the same vehicle over the same route.

Human Operator Optimization.—The discussion of how the human operator adapts has assumed up to now that the criteria governing operator performance remained the same while the operator adapted to changes in the controlled element or the environment. When the operator faces a change in criterion, the problem is somewhat different and may be termed

"optimization" rather than "adaptation." To optimize, the operator must first know what the change in criterion is. In practical terms, this will usually have occurred as a result of some change in the goals of the control process that becomes known to the operator. Manual control is governed by criteria such as speed, accuracy, reliability, economy, safety, and enjoyability in performing a task or function which in turn reflects the mission or goals of the control activity in progress. Changed situations often revise the order of importance of such criteria, as the above driving examples indicate. The operator's source of information is his understanding of the entire control situation and its context. It is only occasionally that this can be reduced to a quantifiable criterion function or figure of merit. Optimization is thus much harder to represent adequately in a human operator model than is adaptation.

ADAPTIVE DESCRIBING FUNCTION MODELS.—When the human operator is represented in a model in terms of parameters of a describing function, e. g., gain, lead, and lag coefficients, operator adaptation is represented by a means of changing such parameters. Gradual changes usually involve an outer control loop which varies parameters so as to optimize some criterion or weighted combination of criteria (figure of merit). An alternative suggested for very fast adaptation involves two or more preestablished sets of parameters/ and switching rules for changing among them.

Criticisms of adaptive models which represent human operator behavior in terms of describing function parameters can begin with the criticisms which apply to stationary describing function models, that is—

(1) By their form these models cannot represent the operator's highly important data reduction function, including the reduction of adaptive information.

(2) They imply that an operator's response at a point in time is a function of input at a particular point in time, failing to represent human memory, planning, and prediction processes.

(3) They fail to incorporate any explicit representation of the task performed.

As a consequence of the last two limitations, adaptation takes on a highly arbitrary character in describing-function models. The model can neither store how the controlled element behaved in the past, nor predict how it is expected to behave in the future. The describing function has no "expectancy," that is, no predicted output with which to compare the output with and judge whether or not adaptation is called for. It has no model of the control situation and possible changes in it to enable it to recognize the nature of a deviation in expectancy and to adapt appropriately to the change. Adaptation is thus nothing more than the blind shifting of parameters according to arbitrary adjustment rules. This is true whether the change to be adapted to is in the controlled element, the environment, or the criterion or "figure of merit" being met.

Investigators who have never employed predictive models to represent the behavior of the operator in stationary tasks have introduced such elements in order to explain how the operator adapts. Elkind, Young, and their colleagues, for example, have posited an internal predictive model to enable the operator to detect and identify changes which require modification of describing function parameters of the model (refs. 8 to 10).

This introduction by Elkind, Young, and their colleagues of a predictive element into their descriptions of the operator's adaptive behavior is a large step in the right direction. However, if the human operator has an internal predictive model that he employs in adapting, is it not likely that he employs the predictive model for control as well as for adaptation? This is to say, is not the Ziebolz controller the appropriate model of the human operator to begin with, whether in adaptive or stationary tasks?

ADAPTIVE ZIEBOLZ CONTROLLER MODELS.—Consider some of the adaptive features of the Ziebolz controller, as illustrated in figures 2 and 3. Adaptation in the Ziebolz controller consists in changing the characteristics of the fast-time model to correspond to changes in the controlled element or the environment. The mechanization of such changes requires that changes be either sensed directly or inferred from the behavior of the controlled element, or both. This is precisely what the human operator must do.

To infer the need for adaptive changes, the human operator compares actual system response with what he has predicted, based on his understanding (model) of the control situation. Since predictions are based on the operator's internal model of the situation, errors in prediction are evidence of errors in the model. Adaptation to inferred changes is the process of detecting errors in prediction, and of adjusting the internal model so as to eliminate them. This same process of manual adaptation could be employed on an adaptive Ziebolz controller model of the human operator. Ziebolz controller adaptation would be effected by adjusting the Ziebolz controller fast-time model so as to minimize errors in prediction. To mechanize this adjustment, an automatic model-matching technique could be employed, changing the parameters of the fast-time model according to the steepest descent or gradient techniques, for instance, to reduce the discrepancy between model-based predictions and behavior of the real controlled element (ref. 11). To compare the responses of the model and that of the real controlled element, they must, of course, both be responding to the same control action. This requires either the application of test signals to the real controlled element, or the storage of control actions used with the real controller for application, in fast time, to the model. The human operator seems to do the latter almost continuously, and the former at times.

Since the fast-time model is employed for control as it is being tested for adaptation, it must be time-shared. For example, the model could be alternated, every other cycle being employed for adaptation rather than control.

The differences between Ziebolz controller adaptation and that of describing-function models are significant:

(1) Adaptive changes are not made in controller parameters, but instead in simulation parameters.

(2) These changes are not arbitrary, but represent changes in the fast-time model that correspond to veridical changes in the controlled element or the environment.

(3) Adaptive changes are based on errors in accelerated time predictions, and corrections to eliminate errors in predictions can be made off-line, iteratively, at high speeds.

(4) The control actions being applied to the real controlled element can be applied to the model off-line without affecting control.

However, the most satisfying feature of the Ziebolz controller model is that it resembles, to quite a reasonable approximation, the process of human operator adaptation, as this was previously described.

OPTIMIZATION.—The human operator infers the need for adaptive change in his internal model of the control situation on the basis of errors in prediction. This inference implies a standard for judging what is and what is not a "good" prediction, that is, a criterion or criteria of performance. All adaptive control systems are governed by one or more criteria of system performance. In most automatic adaptive systems, criteria are represented by a "figure of merit," reflecting the goodness or badness of control over an interval of time in the past according to a weighted combination of criterion signals.

In the case of the human operator, performance assessment may be on a very broad basis, reflecting many different goals served by the system, plus the subgoals of the human operator. The human "figure of merit" may be highly flexible and subjective, provided only

that overall system goals are reached. However, mechanizable adaptive systems, including the Ziebolz controller, require some defined and measured figure of merit that can be built into the controller. Thus it is typical of automatic controllers, in comparison with the human operator, that they function on the basis of much more narrowly defined criteria of performance. Since all control systems, manual and automatic, function in the pursuit of human goals, ultimately only man can exercise judgment as to what it is that the control system is to achieve and when this goal or criterion has changed.

An optimizing as opposed to an adaptive system can be defined as one in which the criterion of performance, rather than the controlled element or the environment, is changed. In such a case, a change in controller response may be called for even though instantaneous input and output, input and output time histories, the controlled element; and the environment are precisely the same. Such changes are routine in human operation. The control of a submarine in wartime will differ considerably according to whether the consideration of (1) time to intercept a target, (2) quietness in the presence of enemy listening devices, (3) precision in holding attitude during weapon delivery, or (4) safe movement through uncharted waters is the predominant criterion governing steering, speed, and depth control.

An important distinction between the Ziebolz controller and other human operator models is reflected in how optimizing changes are made in response to changes in the figure of merit. In describing-function models, model changes in response to all three different classes of change (that is, in the controlled element, environment, or in the criterion or figure of merit) are accomplished in the same basic way, namely, by changing the coefficients of gain, lead, and lag signals or related terms in a controller equation. In the case of the Ziebolz controller, however, entirely different changes are necessary for optimization (adaptation to criterion changes) as opposed to controlled element or environmental changes. Only the latter two can be adapted to by changes in the fast-time analog as described above. Optimization takes place in the Ziebolz controller by a process in which the fast-time analog of the controlled element and its environment stays the same, and, hence, the predicted value of the controlled variable with any given initial conditions and controller response is the same before and after adaptation. What must be adjusted, then, is not the fast-time analog and its prediction, but the way in which the prediction is utilized for control.

To illustrate these differences between adaptation and optimization via Ziebolz controller, consider the automatic control of a space rendezvous via such a controller. Suppose a trial control mode is in effect, with the controller testing possible rendezvous programs on the basis of a predicted fuel consumption criterion and selecting that program tested which accomplished the rendezvous with minimum fuel. Such a controller would adapt by changes in the fast-time model that improve prediction. For example, predictions might be in error because the model had the mass of the spacecraft represented erroneously because less fuel had been consumed than allowed for in the mission profile. Detection of this error and correction of the model would constitute the process of Ziebolz controller adaptation.

Suppose, then, that because of the added fuel, it was no longer necessary to conserve fuel in the rendezvous. The criterion by means of which trials were judged could therefore be switched from minimum fuel to a more appropriate one, for example, minimum time. Then, from the programs tested via trial control, that one which accomplished the rendezvous most quickly could be selected. The change in the criterion employed to evaluate trial control activities would constitute the process of Ziebolz controller optimization.

If the rendezvous were governed by a weighted figure of merit, incorporating time to dock, fuel consumed, and perhaps other factors, any changes in the relative weights would result in different assessment of tested alternatives and would also change Ziebolz control action by "optimizing" rather than by "adaptation."

In the predictive human operator model of figure 2, predictive control is exercised in the outer loop. Changes in "response set" or "compensation" which reduce predicted error according to criteria that are weighted and entered into the evaluator are ever in progress. These ongoing changes are not optimizing per se, but rather are the predictive component of this form of predictive control. Optimizing changes occur when the weighting of criteria ("figure of merit") is changed. The predictive control system will then function to minimize predicted error, now evaluated according to the changed figure of merit.

In figure 3, where predictive control is exercised in the inner loop, adaptation again consists of veridical changes in the fast-time model. Optimization in this case consists in adjustments in how the predictions produced by the model are sampled and processed to form control signals. These adjustments may be made on the basis of fast-time predictions of criterion information. They might consist of changes in prediction span, the gain of the sampled prediction, or other aspects of the sampling and processing of the predictive circuits. In the diagram as shown in figure 3, adjustments in gain and prediction span are entered from an external source. Obviously, these adjustments could be mechanized so they were automatically varied as a function of changes in a figure of merit.

THE HUMAN OPERATOR FUNCTION THAT CANNOT BE MECHANIZED.—The goals of automatic and manual control are, in the final analysis, always human goals. Criteria for evaluating control activity (which are a form of goal) are of necessity ultimately human criteria. The process of adjusting a control system in accord with stated quantifiable criteria is a trivial one compared with the problem of properly selecting and weighting the goals and criteria in the first place, and of changing them as appropriate to changed circumstances.

It is this aspect of optimization, the choosing, weighting, and the appropriate revision of goals or criteria, that can never be adequately represented by a mechanizable human operator model. The mechanism, after all, has no goals of its own; it functions only in the pursuit of human goals. Thus the mechanism cannot be expected to decide what the goals of its human users are, or when and how they have changed. The conception of and choice among goals is the human function that cannot, in principle, be accomplished mechanically.

REFERENCES

1. Sheridan, T. B.: On Precognition and Planning Ahead in Manual Control. Presented at the IEEE Fourth National Symposium on Human Factors in Electronics (Washington, D. C.), May 1963.
2. Ziebolz, H.; and Paynter, H. M.: Possibilities of a Two Time-Scale Computing System for Control and Simulation of Dynamic Systems. Proc. of the National Electronics Conference, vol. 9, Feb. 1954, pp. 215-223.
3. Sheridan, T. B.; Johnson, W. M.; Bell, A. C.; and Kreifeldt, J. G.: Control Models of Creatures Which Look Ahead. Proc. Fifth National Symposium of IEEE on Human Factors in Electronics (San Diego, Calif.), May 5-6, 1964.
4. Sheridan, T. B.: Three Models of Preview Control. Proc. Sixth National Symposium of IEEE on Human Factors in Electronics (Boston, Mass.), May 1965.
5. Kelley, C. R.: A Predictor Model of the Human Operator. Paper presented to the NASA-University of Michigan Working Conference on Manual Control (Ann Arbor, Mich.), Dec. 1964.
6. Kelley, C. R.: Further Research on the Predictor Instrument. Tech. Rep. 252-60-2, Office of Naval Research, Dec. 1960.
7. Kelley, C. R.: Manual Control: Theory and Applications. Office of Naval Research Technical Report, June 1964.

8. Elkind, J. I.; Kelley, J. A.; and Payne, R. A.: Adaptive Characteristics of the Human Controller in Systems Having Complex Dynamics. Proc. of Fifth National Symposium of IEEE on Human Factors in Electronics (San Diego, Calif.), May 5-6, 1964, pp. 143-159.
9. Elkind, J. I.; and Miller, D. C.: Process of Adaptation by the Human Controller. Second Annual NASA-University Conference on Manual Control. SP-128, 1967, pp. 47-63.
10. Young, L. R.; Green, D. M.; Elkind, J. I.; and Kelley, J. A.: Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control. IEEE Trans. on Human Factors in Electronics, vol. 5, no. 2, Sept. 1964, pp. 6-13.
11. Bekey, G. A.; and Humphrey, R. E.: Review of Model Matching Techniques for the Determination of Parameters in Human Pilot Models. Rep. 9865-6002-MU000, Space Technology Labs., Nov. 1962.

IV. COMPUTER PROCESSING OF MANUAL CONTROL RECORDS

12. Relationships Between Fourier and Spectral Analyses

Lawrence W. Taylor, Jr.
Flight Research Center, NASA

N68-15913

About 2-1/2 years ago the Flight Research Center was preparing to analyze human response data for a joint NASA-USAF-Cornell program using a ground based simulator and the variable-stability T-33 airplane. A decision to use expressions of the cross- and power-spectral density functions involving Fourier transforms instead of the cross- and auto-correlation functions led to certain simplifications which raised some questions. These questions in brief were:

(1) Does $\bar{\Phi}_{xy} = F[R_{xy}(\tau)] = \frac{F^*(x)F(y)}{2T}$?

(2) Does $Y_p = \frac{\bar{\Phi}_{ic}}{\bar{\Phi}_{ie}} = \frac{F(c)}{F(e)}$?

(3) Does $\rho^2 = \frac{|\bar{\Phi}_{ie}|^2}{\bar{\Phi}_{ii}\bar{\Phi}_{cc}} = 1$?

where

$$R_{xy}(\tau) = \frac{1}{2T} \int_{-T}^T x(t)y(t+\tau)dt$$

$$F[x(t)] = \int_{-T}^T e^{-j\omega t} x(t)dt$$

These questions were discussed in a paper (ref. 1), entitled "Discussion of Spectral Human-Response Analysis," presented at the Second Annual NASA-University Conference on Manual Control at Cambridge, Mass., on Feb. 28-March 2, 1966. Discussion of the paper from the floor indicated that many persons were of the opinion that Fourier transforms could be used and the simplifications could be made only if the input was the sum of sine waves. In addition, the specific suggestion was made to analyze two samples of a random variable using two measures of the cross-correlation function

$$\bar{\Phi}_{xy_1} = F[R_{xy}(\tau)]$$

and

$$\bar{\Phi}_{xy_2} = \frac{F^*(x)F(y)}{2T}$$

to show that the two are not equivalent.

In response to this suggestion, two independent random signals were produced by using a "random signal generator." Portions of the time histories of these signals are presented in figure 1. The cross-correlation function of the two signals was computed. The results are shown in figure 2.

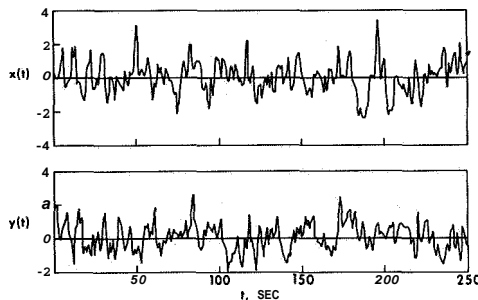


Figure 1.—Time histories of two independent random signals.

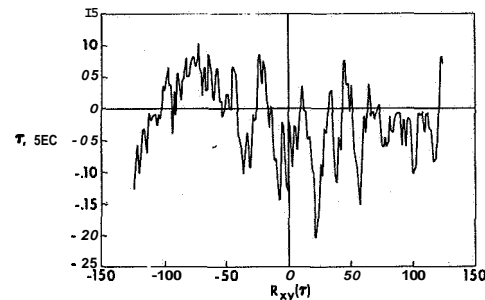


Figure 2.—Cross-correlation function of two random variables.

Finally, the cross-spectral density function was computed using both expressions with identical results (fig. 3). In order to obtain identical results, however, it is necessary (1) not to limit the number of "lags" or the arguments of the cross-correlation function and (2) to use the same "smoothing" operation (if one is used) in both cases. If this is not done, differences will result which can be misinterpreted as differences caused by the lack of equivalence between the two expressions of cross-spectral density. The effects of limiting the number of lags and of smoothing are shown in figures 4 and 5.

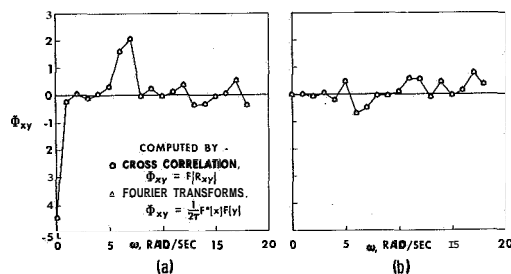


Figure 3.—Comparison of cross spectra computed two different ways. (a) Real parts. (b) Imaginary parts.

This is not to say that some form of smoothing should not be used in the case involving random signals. In paper 9 of this volume the point was made that erroneous results would be obtained if only the raw estimates of the cross spectra were used. The point is that the cross-spectral density function expressed in terms of Fourier transforms can be smoothed just as easily as that expressed in terms of the cross-correlation function. The key question is which computer algorithm is more efficient, since both can be made to do identical operations. The Fourier method often has the advantage of higher speeds and the ability of processing a greater mass of data.

Another factor that, unfortunately, has caused confusion concerning the equivalence of the cross-spectral density expressions was an error in the "proof" presented in reference 1. The limits of integration in taking the Fourier transform of the cross-correlation function were indicated to be $\pm T$ instead of the correct values of $\pm 2T$. This correction has been made in the more detailed proof of equivalence in appendix B of paper 9. Similar proofs can be found in many references, including those listed in reference 1.

It should be noted that when random signals are used as inputs to the compensatory tracking task and some form of smoothing or averaging is required it is no longer possible to

make simplifications that could otherwise be made. This fact is treated in paper 9, especially for the linear-correlation coefficient ρ .

When smoothing is not required ($i = \Sigma$ (sine waves))

$$y = \frac{F(c)}{P \quad Ne}$$

But when smoothing is required (i is random)

$$Y_p = \frac{\sum_{k=-K}^K W_k F_k^*[i(t)] F_k[c(t)]}{\sum_{k=-K}^K W_k F_k^*[i(t)] F_k[e(t)]}$$

where W_k is the weighting function, and

$$\rho^2 = \frac{\left| \sum_{k=-K}^K F_k^*[i(t)] F_k[c(t)] \right|^2}{\sum_{k=-K}^K F_k^*[i(t)] F_k[i(t)] \sum_{k=-K}^K F_k^*[c(t)] F_k[c(t)]}$$

The expected value of ρ^2 is a function not only of the true linear coherence of i and c but also depends on the bandwidth or ensemble over which it is averaged. Figure 6 shows an approximate expression for its expected value as a function of the bandwidth times the record length, or degrees of freedom. The true value of ρ^2 in the example shown is 0.5. Shown in the same figure are a few points that were calculated using experimental data. The experimental points, presumably because of a lack of independence in the data, indicate that a greater number of degrees of freedom would be required than indicated by the theoretical curve to obtain the same accuracy. Since it is not unusual that only a few degrees of freedom be used in computing ρ^2 at low frequencies, the accuracy of such values would be questionable.

It is hoped that the questions concerning the use of Fourier transforms in the analysis of random signals have been resolved and that computational efficiency will become the criterion for selecting which method should be used for a specific application.

CONCLUSIONS

From the preceding discussion, the following conclusions may be drawn:

(1) It is correct to express the cross-spectral density as either (1) the Fourier transform of the cross-correlation function or (2) as the product of Fourier transforms.

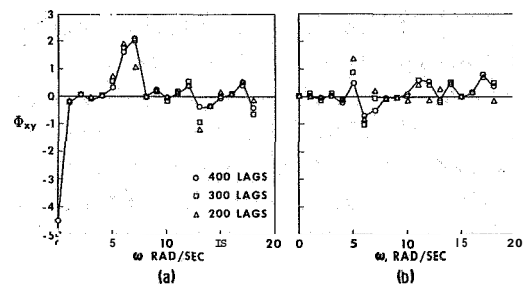


Figure 4.—Effect of number of "lags" on the estimate of cross spectra. (a) Real parts. (b) Imaginary parts.

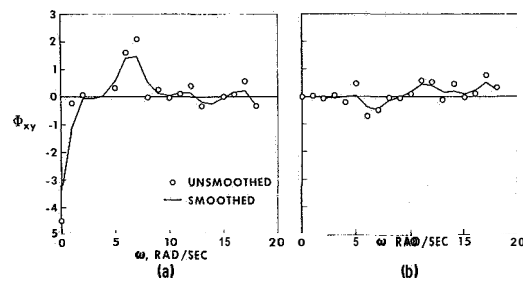


Figure 5.—Effect of smoothing on cross-spectra estimates. (a) Real parts. (b) Imaginary parts.

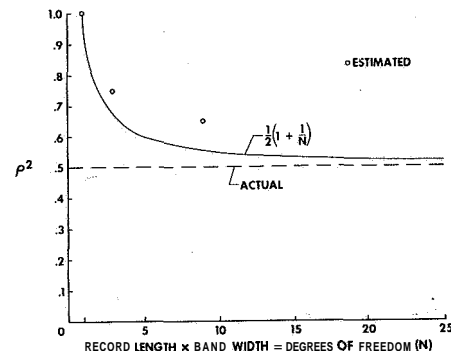


Figure 6.—Bias in linear correlation coefficient estimates.

(2) In order to obtain precisely the same results from the two expressions of cross-spectral density, it is necessary (1) not to limit the number of lags of the cross-correlation function and (2) to use identical smoothing operations in both cases.

(3) When smoothing is not necessary, such as when a sum of sine waves is used as the input, the expressions of Y_p and p can be simplified. When smoothing is necessary, such as when a random signal is used as the input, the expressions for Y_p and p cannot be simplified but can still be expressed in terms of Fourier transforms.

(4) Care must be taken to assure that an adequate number of degrees of freedom are used in computing the linear-correlation coefficient, especially at low frequencies.

REFERENCE

1. Taylor, Lawrence **W., Jr.**: Discussion of Spectral Human-Response Analysis. Second NASA-University Conference on Manual Control, NASA SP-128, 1967, pp. 403-412.

V. CONTROLLED ELEMENTS

13. Manual Control System Performance With Quickened Display, State Variable Display, and Display Gains*

F. A. Muckler and R. W. Obermayer
Bunker-Ramo Corporation

N68

15914

Twelve college subjects performed compensatory tracking with acceleration control dynamics ($0.5s^2$) and a low-frequency forcing function composed of the sum of six sinusoidal components. One group of six subjects performed with a quickened display and the second group of six subjects performed with a state variable display. The state variable display presented system error and system rate error on separate display indices, while the quickened display presented this information combined as the deviations of a single index. All subjects were presented all combinations of two levels of display error gain and three levels of display rate-error gain. Quadratic performance measures were collected to provide information for synthesis of optimal manual control systems under different performance indices. Further, results of a multiple regression analysis, amplitude ratio/phase data, and amplitude distributions are discussed to reflect on display-man-controller performance.

Over the past decade, a change in the conceptual framework of control system theory has been taking place. Rather than design toward a stereotyped concept of a "good" system, the newer approaches deal with optimal system tradeoffs within general performance criteria. This study is an attempt to exercise the philosophy of modern optimal control theory in the synthesis of a manual control system (cf., ref. 1).

Fundamental to modern control theory is the concept of state. The state of a dynamic system may be taken as that set of measurements on the system, which together with a complete description of the forces acting on the system suffice to predict the future behavior of the system. For a system which can be described as a set of linear differential equations, there are as many state variables as the order of the differential equations, and the position, rate, acceleration, and higher order derivatives can constitute a set of state variables.

Also fundamental to optimal control theory is the requirement for the specification of a performance index. While optimal system performance may be in terms of a tradeoff between a number of factors, the theory requires that all performance considerations be collapsed to a mathematical expression defining a single metric. The optimal system, then, is one which maximizes or minimizes this metric.

*This research was sponsored by the National Aeronautics and Space Administration under Contract NAS2-3113.

An example of optimal control theory which is pertinent to the current study is Kalman's solution for optimal linear systems (ref. 2). Kalman's solution applies to linear systems of any order where the performance index is expressed in quadratic form.

In matrix form the system equations and the performance index take the following form:

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} + \mathbf{G}\mathbf{U}$$

$$J = 1/2 \mathbf{X}^T(t_f) \mathbf{S} \mathbf{X}(t_f) + 1/2 \int_0^{t_f} [\mathbf{X}^T \mathbf{Q} \mathbf{X} + \mathbf{U}^T \mathbf{R} \mathbf{U}] dt$$

Here, \mathbf{X} is the state vector, \mathbf{U} the control vector, and J the performance index. The problem is to find a control vector \mathbf{U} , which causes a desired change in the time history of the system state, and which minimizes the performance index.

The performance index is composed of the weighting of the state at terminal time $\mathbf{X}(t_f)$, the time history of the state variables during the intermediate trajectory, and the time history of the use of control. With a quadratic form, the terms are weighted cross products and squares of the state variables; for example,

$$\mathbf{X}^T \mathbf{Q} \mathbf{X} = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12} & Q_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = Q_{11}(X_1)^2 + 2Q_{12}X_1X_2 + Q_{22}(X_2)^2$$

Thus, for example, a term in the quadratic performance index could correspond to the mean-squared-error accumulated while attempting to maintain a given flight trajectory.

For any system of this quite general form, Kalman asserts that the optimal control \mathbf{U} is proportional to each state variable, that is, a linear feedback of the state vector. The optimal gains are given in terms of the solution of a matrix Riccati differential equation (ref. 2); this solution, however, is specific to the regulator control problem and may not yield optimal gains for the servomechanism control problem.

This is a study in the synthesis of manual control systems with acceleration control dynamics. The controlled element, then, is a second-order system which is described with two state variables. The state variables may be taken to be system error e and system rate error \dot{e} . If we exclude considerations of reaching a specific terminal state, the quadratic performance index takes the following form:

$$J = 1/2 \int_0^{t_f} (Q_{11}e^2 + 2Q_{12}e\dot{e} + Q_{22}\dot{e}^2 + R\mathbf{U}^2) dt$$

$$= K_1 \overline{e^2} + K_2 \overline{e\dot{e}} + K_3 \overline{\dot{e}^2} + K_4 \overline{\mathbf{U}^2}$$

The performance index is a given weighting of the mean-squared error $\overline{e^2}$, the mean error-rate-error product $\overline{e\dot{e}}$, and the mean-squared control action $\overline{\mathbf{U}^2}$.

If a linear control action is proportional to each state variable, for example,

$$\mathbf{U} = \mathbf{G}_1 e + \mathbf{G}_2 \dot{e}$$

for a given weighting of terms in the performance index, a specific combination of control gains is required to minimize the performance index. For a manual control system, these

gains depend on the response of the human operator. Unless the human operator fortuitously adopts the proper response, some means must be found to produce the proper human response in order to effect an optimal solution. Fortunately, a model of human response for these controlled element dynamics which usually is a good approximation (ref. 3) contains a term proportional to error and error rate in addition to a reaction-time delay. But some means must be found to alter human response gains if a given optimal solution is to be obtained.

In the current study, an attempt was made to alter human response gains by means of changes in display gain or magnification. Two state variables must be displayed to the human operator, but usually there is some latitude in the manner in which the information is displayed. In particular, system error information may be displayed to the operator so that it appears quite large or quite small. In the case of the second-order controlled element, a display error gain and a display rate-error gain may be selected, corresponding to the state variables of system error and system-rate error. If the human operator responds in some measure proportionately to these changes in display magnification, the manual control system may be tailored to an optimal configuration.

These two items of information may be displayed in different ways. One mode is termed the quickened display. With this display, both error and rate-error signals are added, and displayed as the deviation D of a single bar,

$$D = K_1 e + K_2 \dot{e}$$

Another alternative is to display error and rate-error signals separately as the deviations of two display bars,

$$D_1 = K_1 e \quad D_2 = K_2 \dot{e}$$

We have termed this the state variable display. Suspecting that there may be differences in the effectiveness of influencing human response both displays were included in the current study.

EXPERIMENTAL METHOD

EXPERIMENTAL DESIGN.—In short, the current study is a comparison of quickened and state variable displays, with different display error gains k_1 , and display rate-error gains k_2 . Two levels of display error gain ($k_1 = 0.5$ and 1.0) and three levels of display rate-error gain ($k_2 = 0.5$, 1.0 , and 2.0) were used. Several performance measures were collected, but interest centered on the quadratic performance measures. No changes were made in the controlled element, forcing function, or points of measurement.

Twelve male college students served as subjects in this experiment. All were right-handed and had 20/20 vision, and none had served in a similar study before. Half the subjects tracked with each display, but all subjects received all **six** combinations of two display error gains, and three display rate-error gains. Order of presentation of the experimental conditions was counterbalanced. Since an equal number of subject-order combinations were administered to each experimental display (nested in each display type), but otherwise all subjects received all combinations of the remaining two factors, the data collection procedure satisfied the requirements presented by Winer (ref. 4; p. 319) for a three-factor experiment with repeated measures (case 1).

APPARATUS

Subject's Station.—A CRT display was located directly in front of the subject at a viewing distance of approximately 28 inches. The display face was inclined 5° from the vertical. When the subject was seated with his forearm resting on the incline surface, the spring-centered pencil-type control stick was located at his hand position. The controller was 4 inches long and movable through an arc of 57° . A neon lamp was located directly above the display to warn the subject to get ready for the tracking trial.

The ambient incident illumination was maintained at approximately 4 ft-c through the use of controlled intensity fluorescent lighting. Curtains and partitions were used to screen the subject from distractions, and all of this was contained within another larger closed room. Medium-gray color tones were used throughout in an attempt to produce a restful environment.

CRT Display.—A dual-beam CRT oscilloscope was used as the display device. The 5-inch display face was provided with a standard graticule of centimeter-line etching; a small diamond-shaped symbol was placed in the center of the screen as a standard point of visual fixation. The display could be provided with either one or two moving display elements.

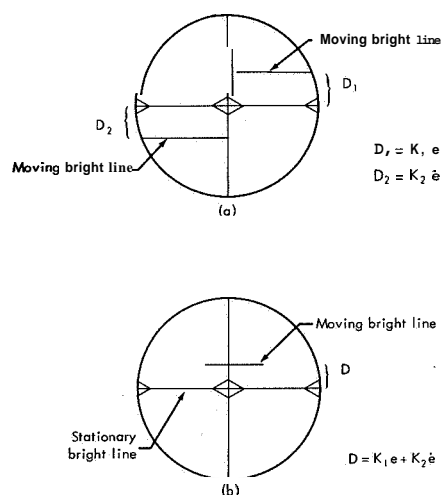


Figure 1.—CRT displays (a) State variable display; (b) Quickened display.

When used as a two-element display (state variable display), a bright bluish-white line was displayed across the left and right sides of the display screen as shown in figure 1(a). Each line extended from a vertical line, dividing left and right sides, to the edge of the screen.

When used as a one-element display (quickened display), the moving element was a centered horizontal line 4-centimeters long; in this case a bright stationary line was provided as a reference, as shown in figure 1(b). In all cases, motion of the display lines was restricted to only up and down directions.

Forcing Function.—If the subjects were to make no response, the only movement on the CRT display face would be due to the input forcing function. The forcing function was identical to that reported by McRuer et al. (ref. 3), constructed from the sum of **six** sinusoidal components of equal amplitudes, random phase relationships, and with the angular frequencies in rad/sec: (1) 0.157, (2) 0.262, (3) 0.393, (4) 0.602, (5) 0.969,

and (6) 1.490. This forcing function has the property that its amplitude is approximately normally distributed; in terms of deviations from the center of the screen, the mean and standard deviation of the forcing function were 0.0 and 1/2 inch.

Analog Computer.—An EAI TR-20 analog computer was utilized to simulate the controlled element ($0.5/s^2$), to generate the quadratic performance scores, and to scale and provide marker signals for magnetic tape recording. The scaling for the control stick (2.5 volts/inch displacement measured along the arc of the top of the control stick) was equal to the display-screen scaling (2.5 volts/inch on the display face). The rate-error signal was compatibly scaled to 2.5 volts/inch/second.

Program Control.—A motor-gear-cam microswitch unit was used to program experimental events. The subjects rested for 14.4 seconds and tracked 105.6 seconds; however, only the last 96 seconds of the tracking time period was scored. Prior to the tracking trial, a warning light was illuminated for approximately 2.4 seconds. The gear system, therefore,

provided a total cam revolution time of 120 seconds. The program was stopped manually for longer rest periods, and a hand-set bell-timer was used to time these break period.

MEASUREMENT

Measurement Equipment.—All measurements other than the quadratic performance measures were derived from signals recorded on a Sanborn seven-channel FM instrumentation magnetic tape recorder. From these tapes the display signals and the subject's control stick response were converted to digital punched tape using a DYMEC 2010B data acquisition system. The analog-digital conversion sampled all signals each 1/2-second, but it sampled the control signal 0.1 second after the display signals. Subsequent measurement and data analysis were performed on the Bunker-Ram0 130 digital computer.

Experimental Measures.—Four classes of experimental measures were collected:

(1) Fourier analysis for each of the forcing function frequencies. Due to the 0.1 second sampling delay, a reaction time delay must be accounted for in these data.

(2) Amplitude distributions of error, rate error, and control signals. Computer counts were made of these signal amplitudes for intervals of 0.1 inch, 0.05 inch, and 0.025 inch.

(3) Multiple regression analysis. During each trial 181 digital samples were collected of error, rate error, and control signals. A multiple regression analysis was performed with the error and rate-error samples used as predictors and the control score as the criterion.

(4) Quadratic performance measures. The following quadratic performance measures were computed on the analog computer and were available at the end of each trial:

Display score (mean-squared error), where e is system error

$$\overline{e^2} = \frac{C_1}{T} \int_0^T e^2 dt$$

Rate score (mean-squared rate error), where \dot{e} is system rate error

$$\overline{\dot{e}^2} = \frac{C_2}{T} \int_0^T (\dot{e})^2 dt$$

Error-rate-error score

$$\overline{e\dot{e}} = \frac{C_3}{T} \int_0^T e\dot{e} dt$$

Control score (mean-squared control deflection), where U is control stick deflection

$$\overline{U^2} = \frac{C_4}{T} \int_0^T U^2 dt$$

where T is the measurement interval time; C_1 , C_2 , C_3 , and C_4 are scaling coefficients for measurement convenience.

PROCEDURE.—Each subject performed for **two** experimental sessions of approximately 2 hours each, one session per day, for a total of 2 days. The inter-session time was approximately 48 hours. On the first day the subjects performed **20** training trials and **two** experimental sessions of 8 trials each. On the second day, the subjects completed the remaining

4 experimental conditions, tracking for a total of 8 trials on each. At the beginning of each day a standard set of instructions was given to the subjects.

A warning light was illuminated for a brief period to indicate that an experimental trial was about to start; the subject then grasped the control stick and was prepared to control when the display started to move. Each trial lasted for 96 seconds, and each trial was separated by an inactive period of 24 seconds. After each fourth trial, the subjects were allowed additional rest periods. The rest periods were administered in a sequence of 3, 5, 3, and 10 minutes, with the sequence then repeating.

EXPERIMENTAL RESULTS

QUADRATIC PERFORMANCE SCORES.—For each of the *six* experimental sessions, eight tracking trials were administered; of these, data from only the last four trials were retained. The first four trials of each experimental session were discarded because these data might be affected by warm-up and transfer effects. Quadratic performance scores were collected for the retained trials and were averaged to provide an estimate of performance with greater stability.

Display Score.—The results of the analysis of variance for the display scores are given in table 1. A significantly less display score results with the quickened display (1.15) than with the state variable display (3.03). The increase in display error gain (0.5 to 1.0) produced a significant decrease in the display scores (2.42 to 1.77). While increases in the display rate-error gain produce increases in the mean display scores for the quickened display and decreases for the state variable display, these differences were not statistically significant.

TABLE 1.—ANALYSIS OF VARIANCE SUMMARY FOR
QUADRATIC PERFORMANCE SCORES

| Source of variation
(a) | F-ratio scores | | | |
|----------------------------|--------------------|--------------------|------------------|--------------------|
| | Display | Rate | Error-error rate | Control |
| Between subjects | | | | |
| A | ^b 12.48 | ^b 18.74 | 1.66 | ^c 6.22 |
| Error A | | | | |
| Within subjects | | | | |
| B | ^c 9.42 | 2.52 | 0.64 | ^b 11.25 |
| AB | 0.07 | 0.47 | 0.47 | 0.33 |
| Error B | | | | |
| C | 0.63 | ^b 24.53 | 0.62 | ^b 36.10 |
| AC | 2.10 | ^c 4.94 | 3.99 | 1.93 |
| Error C | | | | |
| BC | 0.34 | 0.85 | 1.50 | 2.26 |
| ABC | 0.01 | 0.45 | 1.55 | 0.06 |
| Error BC | | | | |

^aKey:

A = displays

B = display error gain

C = display rate-error gain

^bSignificant at 1-percent level

^cSignificant at 5-percent level

Rate Score.—The analysis of variance summary for the rate score is shown in table 1. Increases in display error gain produce an increase in the rate score, but this effect is not statistically significant. The interaction between displays and display rate-error gains is statistically significant. Increases in the display rate-error gain result in decreases in the rate score for both displays, but less with quickened display; that is, the quickened display produces lower scores than the state variable display.

Error-Rate-Error Product Score.—The analysis of variance summary for the error-rate-error product score is shown in table 1. None of these results is significant at the 5-percent level; however, it might be noted that the interaction of displays and display rate-error gains closely approaches significance at the 6-percent level. Smaller error-rate-error product scores result with the state variable display, and, while scores decrease with increases in the display rate gain for the state variable display, the corresponding scores with the quickened display show an increase.

Control Score.—The analysis of variance summary for the control score is given in table 1. The effects of displays, display error gain, and display rate-error gain are statistically significant at the 5-percent level, but none of the interactions is significant:

(1) Significantly less control score resulted with the quickened display (2.72) than for the state variable display (3.40).

(2) Significantly less control score resulted with the lower display error gain (2.88) than for the higher display error gain (3.24).

(3) A Newman-Keuls multiple range test showed that for each increase in display rate-error gain, a significant decrease in the control score resulted (at the 5-percent level).

MULTIPLE REGRESSION ANALYSIS.—The sixth and seventh trial of each experimental condition were converted to digital form by sampling the error, rate-error, and control signals in 1/2-second intervals. Due to the method of sampling and the computer analyses used, the control signal was sampled at approximately 0.1 second after the corresponding error and rate-error samples. With the sampled data (181 of each signal/trial), a multiple regression analysis (ref. 5) was performed using the error and rate-error samples as predictors and the control samples as the criterion. Thus, a linear regression model was fitted to the data which takes the following form:

$$U = B_1 e + B_2 \dot{e} + C$$

While the describing function data (e.g., ref. 3) would lead one to suspect that a linear relationship between control (U) and error (e) could only be approximate, these data allow comparison of the coefficients B_1 and B_2 as to the degree of human-operator dependence on the error and rate-error information as a function of the experimental conditions. In assessing the applicability of this model, it should be noted that overall the data showed a multiple regression coefficient of $R = 0.68$ which suggests that the model accounted for about 46 percent of the variance in the experimental data and that in all the multiple regression analyses (a total of 144) the multiple regression coefficient was significant well beyond the 5-percent level.

Multiple Regression B_1 Weight.—The summary of the analysis of variance for the multiple regression B_1 weight is given in table 2. It will be noted in the variance summary table that the interaction of displays and display error gain and the interaction of displays and display rate-error gain, as well as the respective main effects, are significant at the 1-percent level.

The B_1 weight, the degree of proportionality between the control and error signals, increases with increases in the display error gain. When the display error gain doubles,

TABLE 2.—ANALYSIS OF VARIANCE SUMMARY FOR
MULTIPLE REGRESSION COEFFICIENTS

| Source of variation
(a) | F ratio scores | |
|----------------------------|-----------------------|-----------------------|
| | B ₁ weight | B ₂ weight |
| Between subjects | | |
| A | ^b 10.69 | ^b 11.79 |
| Error A | | |
| Within subjects | | |
| B | ^b 68.35 | ^b 13.37 |
| AB | ^b 13.52 | 0.69 |
| Error B | | |
| C | ^b 17.46 | ^b 16.72 |
| AC | ^b 17.08 | ^c 5.49 |
| Error C | | |
| BC | 0.49 | ^c 4.83 |
| ABC | 0.40 | ^c 3.79 |
| Error BC | | |

^aKey:

A =displays

B =display error gain

C =display rate-error gain

^bSignificant at 1 percent^cSignificant at 5 percent

the B₁ weight nearly doubles for the quickened display, but the effect is less with the state variable display.

The B₁ weight is significantly decreased with increases in the display rate-error gain with the quickened display, but the decreases are not statistically significant with the state variable display.

Multiple Regression B₂ Weight.—The summary analysis of variance results for the multiple regression B₂ weight is shown also in table 2. Here, the highest order interaction of displays, display error gain, and display rate-error gain is significant at the 5-percent level. The B₂ weight; which is a measure of the linear dependence of the control on the rate error signal, increases with increases in display rate-error gain but decreases with increases in display error gain. The B₂ weight is generally greater with the quickened display than with the state variable display.

FOURIER ANALYSIS.—The digital sampled data used for the multiple regression analysis were also subjected to a Fourier analysis. The Fourier analysis was performed by deriving the amplitude and phase of the error signal and the subject's control signal at each of the forcing function frequencies and then computing the amplitude ratio and phase between these two signals. The relationship thus derived would be a describing function of the human operator's response only if he were displayed solely the error signal at unity display error gain. However, this relationship allows assessment of display-operator performance in terms of simple compensatory display performance, that is, when no special attempt is made to display rate-error information to the operator.

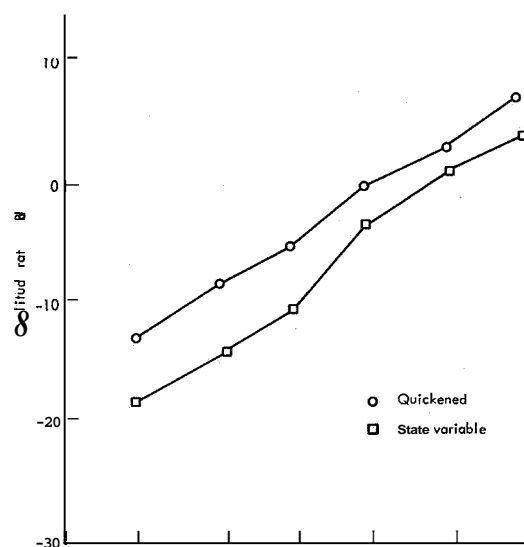
Amplitude Ratio.—The amplitude ratio in decibels plots as a nearly linear relation to frequency on a logarithmic scale. The slope is approximately 20 dB/decade, but it is slightly steeper for the lower display error gains—in particular with the state variable display. While a 6 dB difference might be expected between plots with display error gains of 0.5 and 1.0, the only suggestion of this effect is at the intermediate frequencies with the quickened display. Unfortunately, it is difficult to infer the zero-frequency gain, since the break frequencies are apparently below the range of frequencies in the forcing function.

Phase.—With the quickened display, the only noticeable differences in phase occur at the high and low frequencies of the forcing-function range. Greater lead at the higher frequencies and greater lag at the lower frequencies accompany increases in the display rate-error gain. Increasing display rate-error gain also corresponds to greater lag at the lower frequencies for the state variable display at the low display error gain, but little difference exists for the higher frequencies. For the high display error gain, very erratic phase performance results with the state variable display. The cross-over frequency for these conditions occurs at approximately 0.4 to 0.6 rad/sec.

Quickened versus State Variable Display.—To allow convenient comparison, the amplitude ratio and phase data for the quickened display and the state variable display are shown in figures 2 and 3 averaged over all other conditions. The amplitude ratio is between 2 and 5 dB greater for the quickened display than for the state variable display. Greater phase lead also results with the quickened display. For much of the frequency range, the phase difference between the two displays could be reduced if one of the phase curves were shifted by approximately 0.1 rad/sec, suggesting a lower break-frequency for the quickened display.

It should be remembered, however, that data are for these two displays as collected with two different groups of six subjects. It is not possible to discriminate clearly whether the observed effects are due to display differences or group differences.

Amplitude Distributions.—Based on the digital sampled data, amplitude distributions of the error, rate error, and control signals were computed by counting the number of occurrences within small intervals (as measured along the arc of the top of the control stick, or as displayed on the CRT with unity display error gain, or display rate-error gain). Distributions were computed for each combination of experimental conditions; however, except for different variances as reflected in the display score, rate score, or control score, no striking differences were noted. While normal-shaped distributions resulted with the error and rate-error signals, the control distributions are very flat. This is in contrast with the peaked and tri-modal appearing, distributions found elsewhere, but with a rate control task (ref. 6).



DISCUSSION

MANUAL CONTROL SYSTEM QUADRATIC PERFORMANCE

Computed Quadratic Performance Scores.—If for computational convenience we assume that the human operator responds proportionately to both displayed system error and system

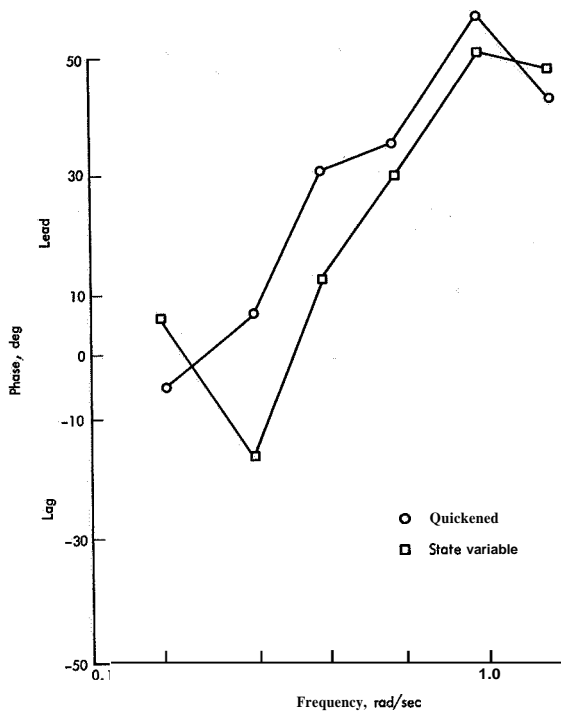


Figure 3.—Phase for quickened versus state variable displays.

ables (e, \dot{e}), as required by the Kalman solution for optimal linear systems, if display gain were effective in changing the feedback gain to each of the state variables. If these results would hold with a human operator in the control loop, the manual control system could be adjusted to optimal performance using display gain to produce the feedback required by the Kalman solution.

Experimental Quadratic Performance Scores.—The experimental quadratic performance scores, averaged over both information displays, are shown in figure 5. A number of differences were noted in comparison with the computed scores:

(1) While mean-squared-error decreases with increases in display error gain, mean-squared-error does not monotonically decrease with display rate-error gain as with the computed scores. High display rate error gain causes a deterioration in mean-squared-error performance (these differences do not reach statistical significance, however).

(2) Mean-squared-rate-error decreases with increases in display rate gain, but deteriorates with increases in display error gain (again, while the effect is consistent, these differences are not statistically significant at the 5-percent level).

(3) The amount of control used, mean-squared-control-deflection, increases with display error gain and decreases with increases in display rate error gain, as do the computed scores.

rate error and is invariant in his response as display gains are changed, the results shown in figure 4 may be computed. This figure shows the quadratic performance scores which would result as display error gain and display rate-error gain are varied. Under the assumption, doubling the deviations seen on the display results in doubled human response, and, hence, doubled open-loop gain. The curves in figure 4 are based on the specific controlled element and forcing function used in this study.

It may be seen from figure 4 that with such a mechanism, for loop gains greater than unity: (1) increasing either display error gain or display rate-error gain causes reduced mean-squared error; (2) similarly, increasing either display error gain or display rate-error gain causes reduced mean-squared-rate-error; and (3) increasing display rate error gain reduces the amount of control (mean-squared-control-deflection), but increases in display error gain increase the amount of control. The results for loop gains less than unity are also shown.

These are the results which would be obtained with a linear feedback of the state vari-

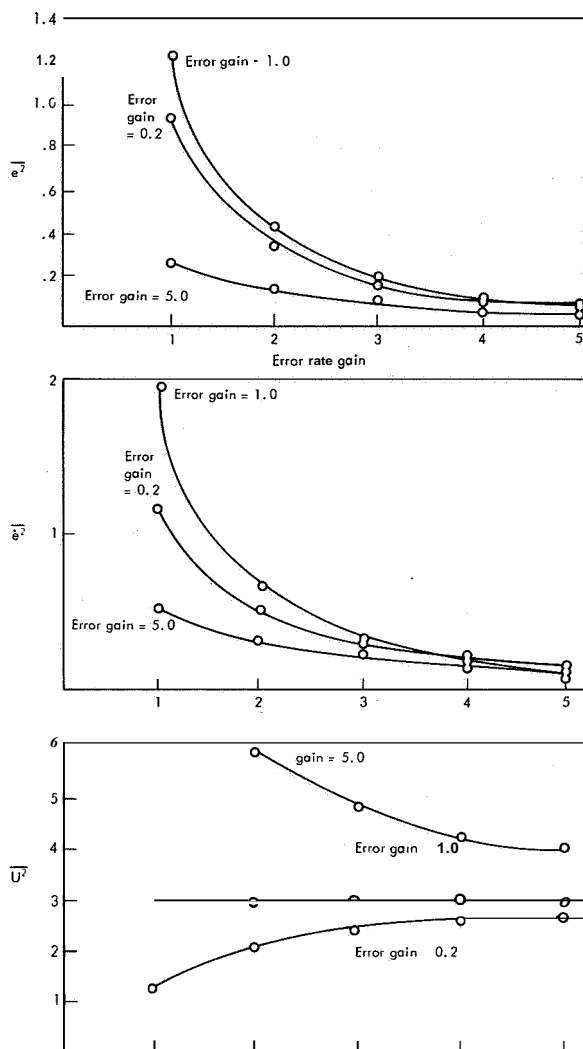


Figure 4.—Computed quadratic performance measures.

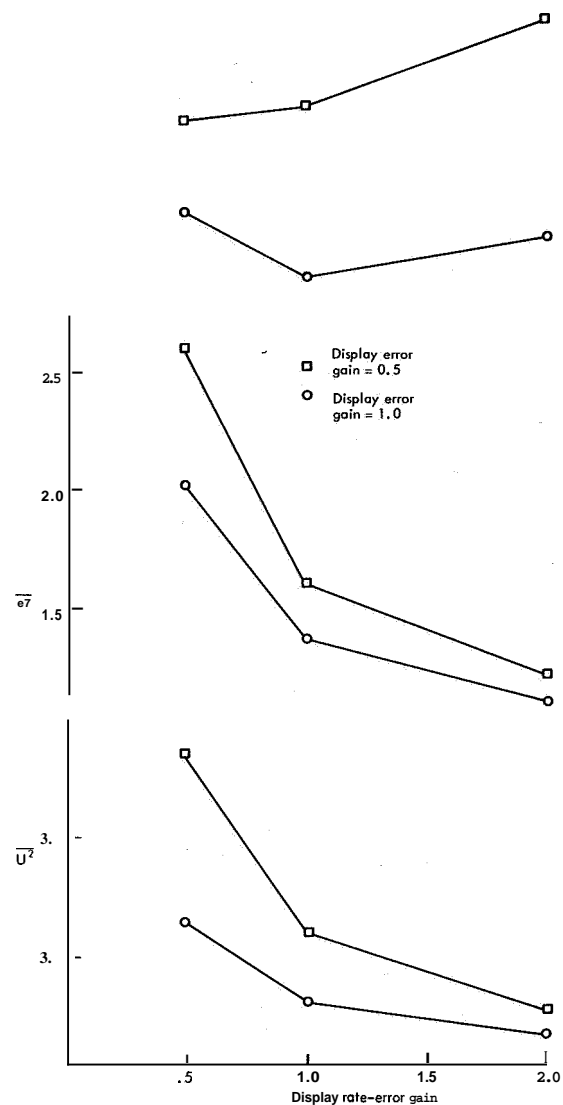


Figure 5.—Experimental quadratic performance measures.

It may be seen that display gain for a specific state variable apparently directly affects the system gain for that state variable—increased gain decreases system error for that state variable—but further has the effect of deteriorating performance on the other state variable. The effect on the amount of control is as if changes in display gain have a corresponding effect on the open-loop gain.

Optimal Manual Control System Performance.—For optimal quadratic performance, a number of tradeoffs must be made in selecting display gains for a second-order manual control system. Increases in display error gain improve mean squared error but increase mean-squared-rate error and vice versa. Both display gains affect the amount of control action used—one increasing the amount, the other decreasing the amount of control.

The proper tradeoff depends on the relative importance of each system error and the relative importance of control action. The approach of optimal control theory is to define a performance-index weighting each system attribute. The theory assumes that these weights are given, and it is up to the design engineer to determine them. The specific definition of

optimality is provided by the definition of the performance index, and, of course, the optimal tradeoff is critically dependent on this criteria1 definition.

To demonstrate the optimal tradeoff as a function of the performance index, table 3 shows the performance index for the experimental data of this study. The quantities indicated by a superscript "a" are optimal table values. Three different results occur with each of three different weightings in the performance index. The weighting of mean-squared-error-rate-error product is arbitrarily taken as zero. When mean-squared-error is weighted heavily, the optimal combination is display error = 1.0, and display rate error = 1.0. When the quadratic performance scores are weighted equally, the highest display gains are optimal. When control is weighted heavily, the optimal combination is display error gain = 0.5 and display rate-error gain = 2.0.

TABLE 3.—PERFORMANCE INDEX WITH DIFFERENT WEIGHTING

| Different weighting | Display error gain | Display rate-error gain | | |
|--|--------------------|-------------------------|-------------------|-------------------|
| | | 0.5 | 1.0 | 2.0 |
| $J = 10\overline{e^2} + \overline{\dot{e}^2} + \overline{U^2}$ | 0.5 | 27.9 | 27.2 | 30.5 |
| | 1.0 | 25.9 | ^a 21.0 | 21.6 |
| $J = \overline{e^2} + \overline{\dot{e}^2} + \overline{U^2}$ | 0.5 | 7.5 | 6.5 | 6.5 |
| | 1.0 | 8.4 | 6.4 | ^a 5.8 |
| $J = \overline{e^2} + \overline{\dot{e}^2} + 10\overline{U^2}$ | 0.5 | 35.8 | 31.7 | ^a 30.5 |
| | 1.0 | 43.1 | 34.4 | 30.7 |

^aOptimal table values

Optimal Quickening Coefficient.—It is clear then that the manual control system is not best for any given combination of display gains, but that the optimal manual control system is a function of the definition of the performance index. In the literature on quickened displays (cf., refs. 7 and 8) some speculation is given on the best combinations of derivative information, the quickening coefficient. It is clear in the context of the preceding discussion that the optimal amount and proportions of display derivative information depend on the nature of desired system performance.

DISPLAY-MAN-CONTROLLER PERFORMANCE

Effect of Display Gain.—Based on the quadratic performance measures, the multiple regression coefficients, and the Fourier Analysis, varying display gain has a corresponding effect on the gain of the display-man-controller combination. As display gain is increased, the appropriate system performance improvement occurs, the average linear dependence of control deflection on display deviation (multiple regression coefficient) increases, and generally the appropriate changes in the Bode plots result. In short, over the range of gains in this study, increased display gain resulted in higher gain for the display-man-controller

combination. However, while no data are presented here on human response to display deviations (only to system error) since increases in display gain resulted in less than proportional increases in the gain of the display-man-controller combination, it can be inferred that increased display gain resulted in decreased human response gain.

Quickened versus State Variable Display.—Two types of display were used in this study which ostensibly provide the same information: the algebraic sum of what is presented with two display bars in one case is displayed as the deviation of one display bar in the other case. However, on most counts there is a striking difference in favor of the quickened display:

(1) Lower display scores, rate scores, and control scores result with the quickened display.

(2) The multiple regression B_1 and B_2 weights are nearly proportional to the display error gain and display rate-error gain, respectively, with the quickened display, while only a fraction of the same change is noted with the state variable display.

(3) Greater amplitude ratio and greater phase lead result with the quickened display. Low phase margin resulted with the state variable display and the high-display error condition.

Nevertheless, given specific system requirements, the state variable display could be preferred to the quickened display. While the quickened display does provide a consolidated display with no required division of operator attention, the state variable display allows the human operator separate assessment of performance with each state variable and does not allow the motions of one state variable to mask the motions of the other. In this study, the effects of display error gain and display rate-error gain were relatively less cross-coupled with the state variable display than with the quickened display. Further, the state variable display elicited a smaller mean-error-rate-error product than the quickened display. Had that been the prime system measure requirement, the state variable display would have been superior.

Under the approach of modern control, the phrase "optimal system performance" has precise meaning only in terms of a specific performance criterion. No matter how simple or complex the control system, the optimal performance criterion must be stated before a valid design decision can be made for human operator performance and his associated display and control devices.

REFERENCES

1. Obermayer, R. W.; and Muckler, F. A.: Modern Control System Theory and Human Control Functions. NASA CR-256, 1965.
2. Obermayer, R. W.; and Muckler, F. A.: On the Inverse Optimal Control Problem in Manual Control Systems. NASA CR-208, 1965.
3. McRuer, D.; Graham, D.; Krendel, E.; and Reiser, W.: Human Pilot Dynamics in Compensatory Systems: Theory, Models, and Experiments With Controlled Element and Forcing Function Variations. AFFDL-TDR-65-15, July 1965.
4. Winer, B. J.: Statistical Principles in Experimental Design. McGraw-Hill Book Co., Inc., 1962.
5. Cooley, W. W.; and Lohnes, P. R.: Multivariate Procedures for the Behavioral Sciences. John Wiley and Sons, Inc., 1962.
6. Obermayer, R. W.; Webster, R. W.; and Muckler, F. A.: The Effect of Four Performance Criteria on Compensatory Rate-Control Tracking. Bunker-Ramo Corp., Canoga Park, Calif., Jan. 15, 1966.

7. Bailey, A. W.; and Sweeney, J. S.: Preliminary Study of Helicopter Attitude Display Systems. NRL Rep. no. 451, 1955.
8. Ely, J. H.; and Bowen, H. M.: Man-Machine Dynamics. Human Engineering Guide to Equipment Design, ch. 5, C. T. Morgan, J. S. **Cook**, and A. Chapanis, eds., McGraw-Hill Book Co., Inc., 1963, p. 242.

14. Sensory-Motor Aspects of Manual Control of High-Inertia Tracking Systems

Russell L. Smith and John Lyman
University of California, Los Angeles

68-15915

A series of experiments were conducted on a high-inertia manual-tracking device (converted 40-mm gun mount) that simulated closely the characteristics of field tracking systems. The principle goals of the studies were to determine the extent to which results derived from research using low- and negligible-inertia tracking simulators could be generalized to high-inertia simulators, and to derive optimum display and control configurations for tracking realistic target trajectories. Important theoretical and practical implications were suggested by the experimental results. It was found that tracking performance was facilitated by increasing display magnification, optical magnification (reducing subject-to-display distance), and proprioceptive cues derived from azimuth and elevation movements of the simulator. Reduction in the display field of view and/or addition of displacement aiding velocity control dynamics also significantly improved performance. Contrary to previous research, no differences were observed in performance with pressure and with movable-type controllers for experienced trackers, although the former controller appeared superior during early learning trials.

Certain manual control tasks, such as tracking surface-to-air and air-to-air missiles, require the use of high-inertia systems. Such systems are high in mass because they carry heavy photographic and optical equipment as well as the operator. Typically, there has been a paucity of experimental studies of high-inertia systems; usually a cathode ray tube or similar device with low or negligible inertia is employed. These latter devices characteristically present operators with the task of tracking a dot undergoing periodic changes through a zero-error position. The forcing functions used differ from missile trajectories in that they may be periodic or random appearing rather than simulating the flight of a preprogramed or continuously programed object, and permanent target losses are usually impossible regardless of how poor an operator may be performing.

A major problem confronting the designer of missile tracking systems is the determination of the extent to which he can apply contemporary findings to his particular needs. A primary objective of the experiments reported here was to provide data relevant to that problem. A second objective was to gain further knowledge concerning more precise explanations of the effects on performance of magnification, field of view, and various unique controller configurations.

APPROACH

Heavy-inertia tracking systems have a number of inherent characteristics which cannot be adequately simulated in low-inertia devices. In order to generalize experimental findings to the field situation with a high degree of confidence, it therefore seemed essential to employ a simulator which resembled the field system as closely as possible. The NOTS-UCLA tracking simulator was developed to fulfill that requirement. It consists of a converted 40-mm

Naval gun mount placed at the center of a white quartersphere screen having a 10-foot radius. With the aid of binoculars on the mount or closed-circuit TV in a remote isolation booth, an operator tracks a moving point of light projected onto the quartersphere surface. Controller movements produce dc voltage variations that provide an input signal for an analog computer that controls the type and proportion of aiding used to drive the mount. Integrated error scores are derived from a sensitive photo-cell on the mount which samples tracking error 24 times per second.

The two trajectories (T-1 and T-2) used in all experiments are shown in figure 1. Although both are ballistic in nature, T-2 provides considerable variation in velocity, acceleration, and course from a true ballistic path. These trajectories were selected to contain representative characteristics of simple and more complex trajectory types found at the Naval Ordnance Test Station.

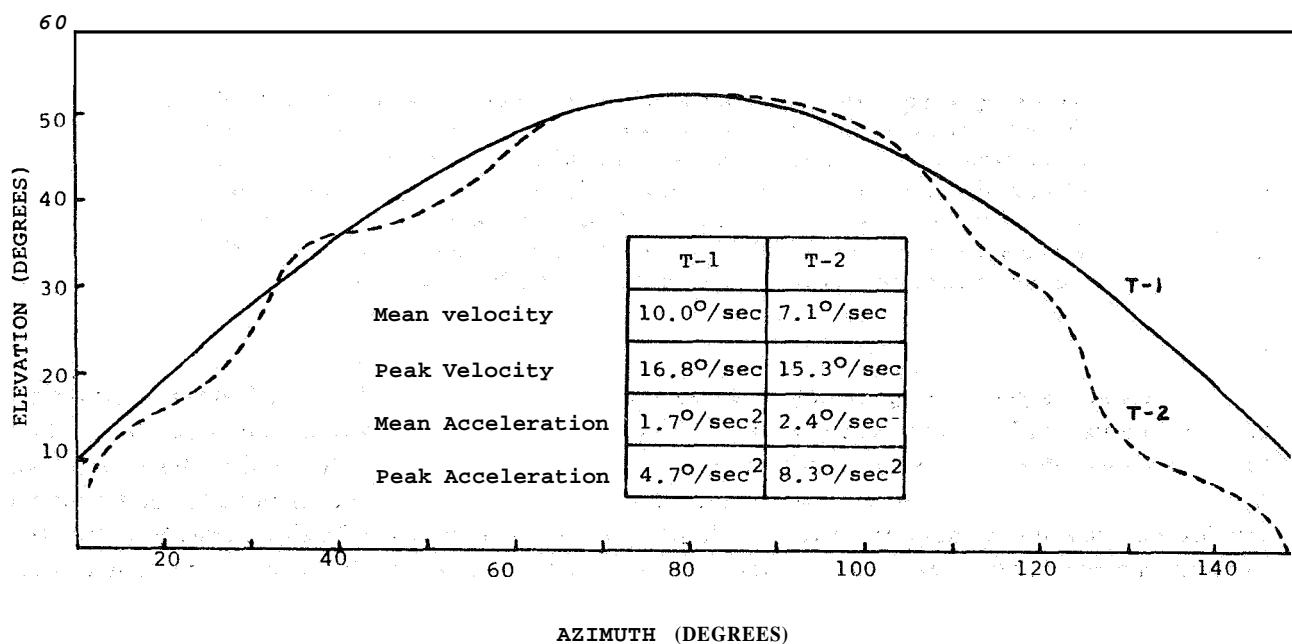


Figure 1.—Spatial and dynamic characteristics of trajectories T-1 and T-2.

Six and four highly trained college students were used in the first and second experiments, respectively, while four experienced and four naive students were employed in the third. Ten data points were obtained from each subject on each condition.

A more detailed account of the following three studies is presented in references 1 to 4.

EXPERIMENT I

PROBLEM.—To identify important interactions among major variables, the initial experiment was designed to include several variables known to have significant effects on performance. It was planned that if the interactions found had little or no practical importance, subsequent studies could concentrate on determining optimum levels of single variables.

Previous studies have shown that tracking performance is facilitated by magnifying tracking error (refs. 5 to 9), by the addition of an aiding component to rate control dynamics (refs. 9 to 12), and by the addition of proprioceptive cues derived from a moving tracking

mount (ref. 12). These and one other variable, trajectory complexity, were included in a single experiment. Specific levels of each variable were the following:

- (1) Proprioception
 - (a) Azimuth and elevation cues (AEP)—The operator rotated in azimuth and tilted his head with the motion of the elevation carriage containing the binoculars.
 - (b) Azimuth only (AP)—Elevation cues were eliminated by replacing the binocular with a TV display fixed to the azimuth carriage.
 - (c) No proprioception (NP)—The simulator was remotely controlled via the TV display.
- (2) Display magnification
 - (a) No magnification, 13' field of view (1x-13")
 - (b) 5x-5°
- (3) Trajectory
 - (a) Smooth (T-1)
 - (b) Irregular (T-2)
- (4) Control dynamics
 - (a) Velocity (time constant=0.0 sec)
 - (b) Velocity plus displacement aiding (0.1 sec)

Levels of each of those variables represented relatively extreme points in our system and were selected on the basis of preliminary studies.

RESULTS AND DISCUSSION.—Figure 2 summarizes the major results of this study. An analysis of variance indicated that all main effects and one first-order interaction, proprioception by trajectory, were significant sources of variation. The single interaction suggests that proprioception facilitated performance more on the irregular trajectory than on the simple trajectory. It can be seen in figure 2 that for each proprioceptive condition, error magnification and displacement aiding independently reduced error scores substantially, although their effects were greater on T-2 than T-1.

The condition employing whole body proprioceptive cues (AEP), magnification, and displacement aiding resulted in outstanding performance on both trajectories. The effect of that condition on T-2 is particularly dramatic in that performance was not significantly different from five of the 12 conditions on T-1 (circled dots).

Although the results of this experiment generally confirmed and extended the findings of previous studies, there is at least one point of departure and another left unresolved. The classic study of Chernikoff and Taylor (ref. 11) demonstrated a significant interaction between control dynamics and trajectory complexity. Velocity control alone and displacement control alone were found best for simple and complex trajectories, respectively, while velocity plus displacement control (time constant = 0.5 sec) was found best for intermediate types of trajectories. Although those findings are generally regarded as valid for low-inertia tracking systems, the results presented herein suggest that for high-inertia systems, the combination of velocity plus displacement dynamics will facilitate performance on most, if not all, realistic missile trajectories.

A confounding variable in this and previous experiments investigating performance effects of error magnification is the visual field of view, i.e., for most lens systems an increase in magnification results in a reduction in field of view. Such was the case in the present study and it cannot be assumed on intuitive grounds that field of view plays an insignificant role in determining performance level. Experiment II was therefore designed in part to investigate this problem empirically.

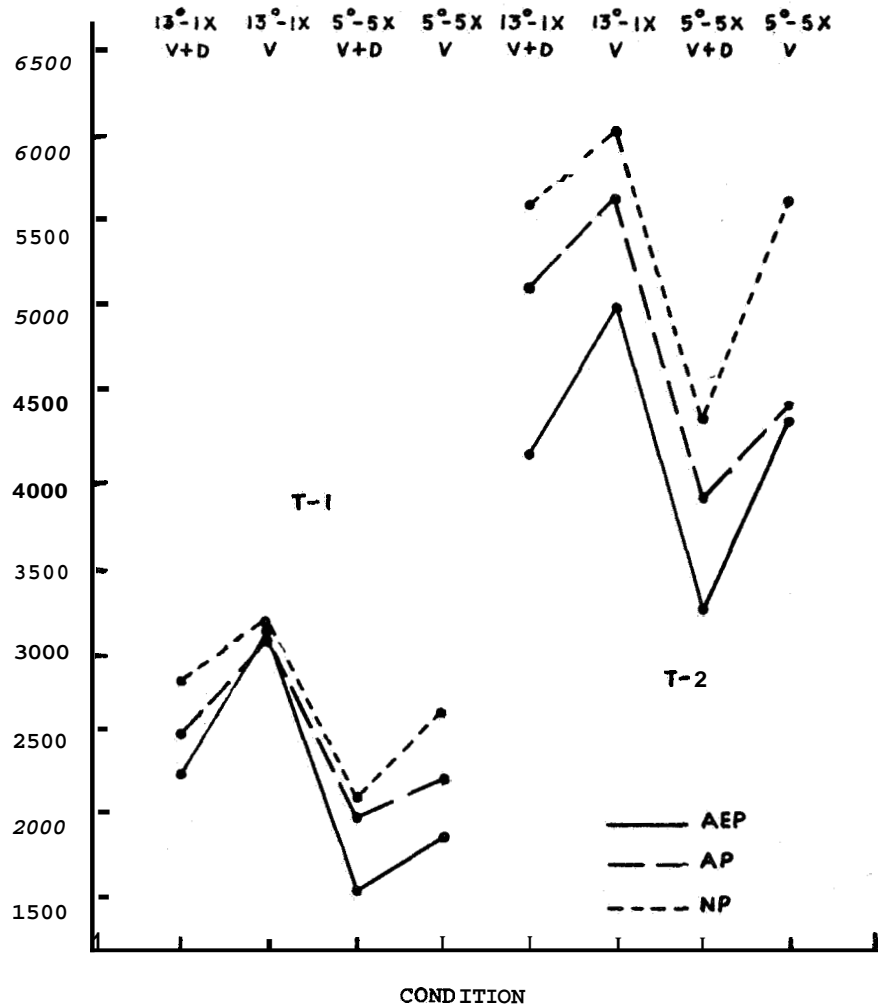


Figure 2.—Tracking error as a function of viewing mode, trajectory, optical system, and control dynamics.

EXPERIMENT II

PROBLEM.—As noted, it apparently has been assumed that magnification per se, rather than the concomitant reduction in field of view, improves tracking performance. It has been suggested that the rim of the visual display acts as a frame of reference for subjectively measuring target movements (refs. 8 and 13). This implies that target movements can be better estimated and thus tracked with magnification since they more closely approach the frame of reference. However, magnification not only results in higher display-target velocities, it also results in a decrease in field of view. Lockard (ref. 14) theorized that field reduction ought to be detrimental to performance since position and velocity of the target cannot be adequately perceived. Evidence supporting that view was presented by Gottsdanker (ref. 15). If the above explanations are correct it would appear that the detrimental effects of a smaller field are outweighed by the facilitative effects of magnification.

A third line of reasoning was given by the authors (ref. 16) and a pilot study provided empirical support. It was suggested that performance is facilitated by magnification not because target motions are more clearly discriminated but because the smaller field of view

forces the operator to react faster. We thus hypothesized that magnification per se has little effect on performance, providing the target is easily observable.

The effect of "forcing the operator to react faster" should be obtainable by yet another method of increasing magnification. Reducing the distance between the operator and display (O-D distance) has been termed "optical magnification" which, although producing the identical effect as display magnification in terms of the retinal angles subtended, does not alter target dynamics as does display magnification. The psychological effect may be similar, however, since increases in retinal magnifications result in apparent increases in target velocities and accelerations. These apparent increases may again result in a reduction in operator reaction time.

For this experiment three interrelated variables were investigated using the off-the-mount TV tracking mode. A zoom lens attached to the simulator provided display magnification variations while masks of different diameters were used to vary the size of the field of view. The third variable, optical magnification, was varied simply by changing the O-D distance. Levels of each variable were the following:

- (1) Field of view (O-D = 72 in.)
 - (a) 1x-15° (1x)
 - (b) 1x-10° (1x)
 - (c) 1x-7.5° (1x)
 - (d) 1x-5° (1x)
 - (e) 1x-3° (1x)
- (2) Display magnification (O-D = 72 in.)
 - (a) 2x-7.5° (2x)
 - (b) 3x-5° (3x)
 - (c) 3.75x-4° (3.75x)
 - (d) 4.1x-3.67° (4.1x)
- (3) Optical magnification (O-D = 30 in.)
 - (a) 1x-12° (2x)
 - (b) 2x-6° (4x)
 - (c) 2.5x-4.8° (5x)
 - (d) 3.4x-3.53° (6.8x)
 - (e) 4.1x-2.92° (8.2x)

The values given in parentheses are the retinal magnifications, reflecting the O-D variations.

RESULTS AND DISCUSSION.—The performance curves of the three sets of conditions on T-1 and T-2 as a function of field of view and magnification are shown in figure 3. The display and retinal (parenthetical values) magnifications presented by the means of T-2 curves have been omitted on T-1 for clarity.

A multiple comparisons test of the means indicated no significant differences between the performances of operators subjected to display magnification or simple field reductions although there is an apparent superiority of the display magnification performance curve on T-1. In both sets of conditions reduction in field of view, down to 3° to 4°, resulted in a significant improvement in performance. These data suggest that the facilitative effect of display magnification is totally due to the concomitant field reduction.

Doubling the optical magnification resulted in better performance than that obtained in the above modes when no display magnification was used. However, subsequent display magnification had no effect on performance regardless of trajectory. It thus appears that operators were reacting maximally under the initial optical magnification so that additional adjustments in display characteristics could not result in further tracking improvement.

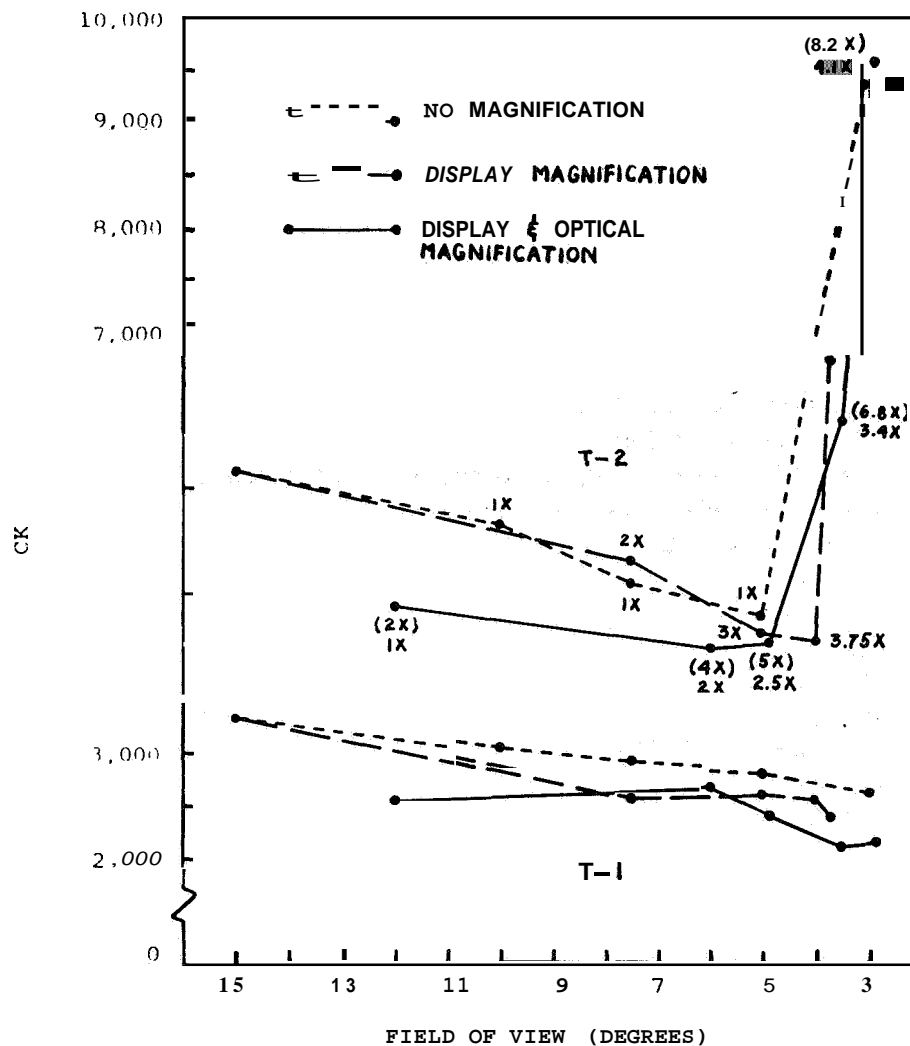


Figure 3.—Tracking error on trajectories T-1 and T-2 as a function of type of magnification and field of view.

Although asymptotes were not yet achieved on the simple trajectory, it is interesting to note that optimum performance on T-2 occurred with a field of view of 4° to 5° and the type or degree of magnification was of negligible importance.

The results of this study generally confirmed our predictions, but we did not anticipate the optical magnification performance curve to deviate significantly from the other curves. It is our present belief that this finding may be explainable in terms of the differences in task difficulty. As the field of view is reduced, the requirement to maintain the target within the remaining display area becomes more and more stringent. Increases in display magnification result in the additional factor, augmented target dynamics. On the other hand, optical magnification only appears to contain these constraints. It therefore seems reasonable to expect higher performance on a task which is inherently easier but appears to be equally difficult.

For all practical purposes optical magnification has little or no utility. It cannot be manipulated in typical tracking situations where optical systems consist of binoculars and it is unlikely that O-D distances would ordinarily be much greater than 30 inches when remote TV tracking is employed. Findings regarding the independent effects of display magnification

and field of view, however, are of both theoretical and practical importance and, if verified, should be considered in the design of optimum field tracking optical systems.

EXPERIMENT III

PROBLEM.—The literature shows considerable differences of opinion about the relative merits of contemporary manual controllers. Principal issues are related to control mobility, pressure versus displacement, and null centering force. A number of studies suggest that if the tracking task is discrete or continuous, and pursuit or compensatory, relatively large control movements may be more favorable than small movements (for example, refs. 8, 17, and 18). The cited experiments used position, rather than velocity dynamics, however, and opposite conclusions were reached in reference 19 when velocity dynamics were used.

In reference 20, it was theorized that the movable controller ought to be superior to the pressure controller on the grounds that the operator receives kinesthetic information of limb movement as well as a certain amount of pressure information from the former while only pressure cues are obtained from the latter. On the other hand, available empirical evidence appears contradictory (refs. 12, 21, and 22). Unfortunately, this evidence is somewhat inconclusive since in reference 21 a true pressure controller, was not used (that is, it had a "slight" movement capability), and the operators in reference 12 had had years of prior practice with a movable control, but only 2 days to adjust to the pressure control.

While the importance of having a strong null centering force on a controller is fairly well established (ref. 23), its importance relative to the variables previously discussed is uncertain.

Our experiment III was designed to assess the merits of the three control variables with inexperienced operators and operators having had considerable experience with a movable, strong centering controller. All operators were given preliminary training on each controller (about 120 runs per controller) followed by eight trials in which data were acquired. The tracking mode was on-the-mount with the binocular display. Table 1 lists the characteristics of the four controllers employed in this study.

| Controller | Mobility | Centering | Transducer |
|------------|------------------|------------------|-------------|
| HM-SC | Highly movable | Strong centering | Resistance |
| HM-WC | Highly movable | Weak centering | Resistance |
| SM-SC | Slightly movable | Strong centering | Resistance |
| NM-SC | Not movable | Strong centering | Capacitance |

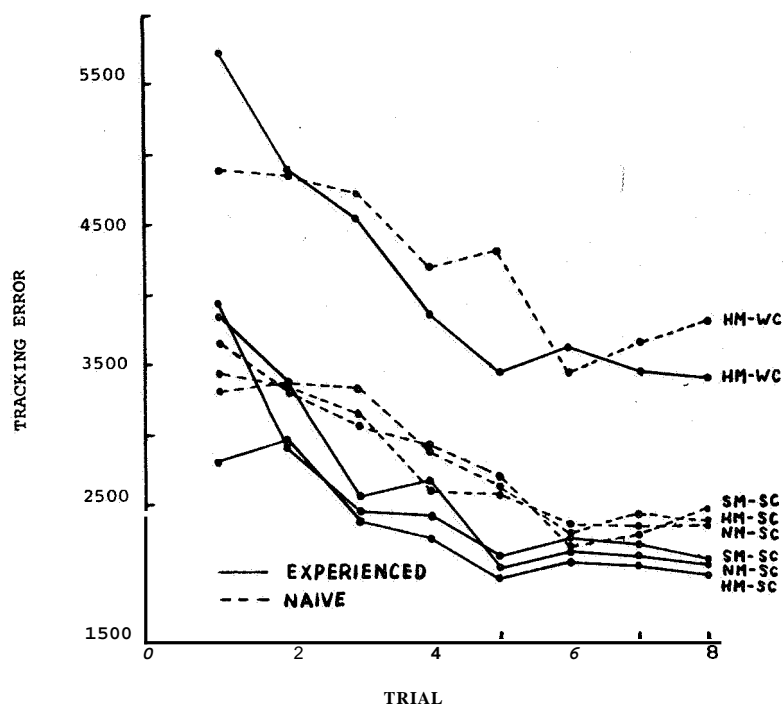


Figure 4.—Learning curves of experienced and naive subjects on all controllers and trajectory T-1.

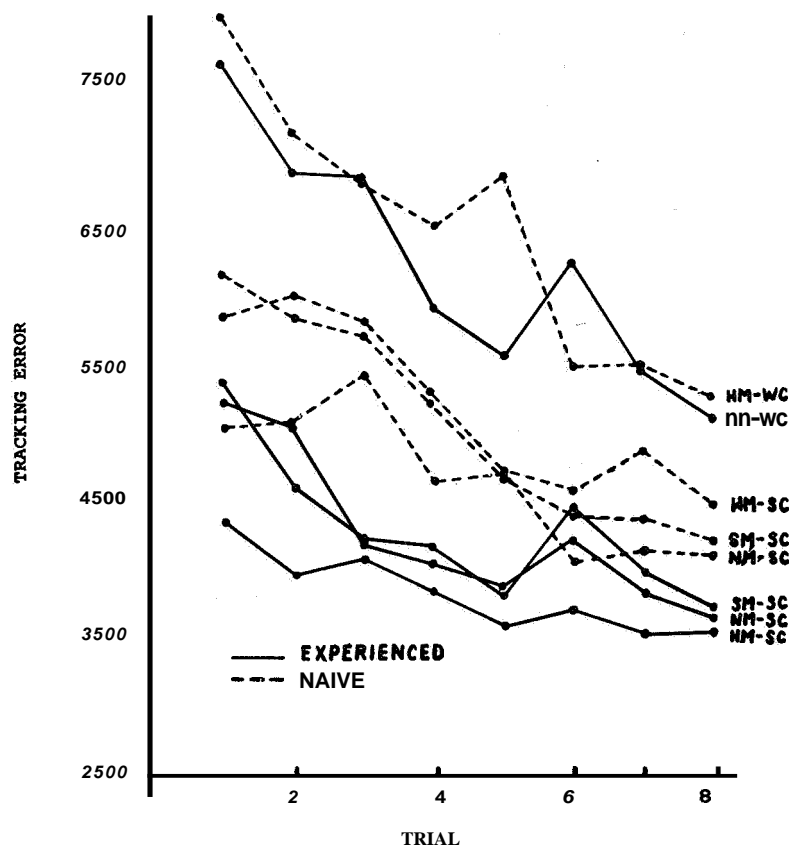


Figure 5.—Learning curves of experienced and naive subjects all controllers and trajectory T-2.

Clearly the finding of most significance was the poor showing of the weak centering control (HM-WC). Based on our results, we can only conclude that strength of centering is the **only** variable of those investigated which was of importance. The initial advantage of the pressure controller with naive operators (fig. 5) was at most short lived.

CONCLUDING REMARKS

Experiments on the high-inertia tracking system suggest that results obtained on low-inertia systems probably may be only partially generalized to field tracking situations. The implication from the results is that the manual tracking situation may involve several specific variables which have not been fully exploited in formal models of tracking behavior.

REFERENCES

1. Garfinkle, D. R.; Smith, R. L.; Groth, H.; and Lyman, J.: Performance Studies on the NOTS-UCLA Tracking Simulator: Effects of Error Magnification, Field of View, Proprioceptive Cues, and Control Dynamics. Rep. no. 63-41, Dept. Eng., Univ. Calif., Los Angeles, 1963.
2. Smith, R. L.; Garfinkle, D. R.; Groth, H.; and Lyman, J.: Effects of Error Magnification, Proprioceptive Cues, Trajectory Characteristics, and Control Dynamics on Compensatory Tracking. Human Factors, 1966.
3. Smith, R. L.; Garfinkle, D. R.; Groth, H.; and Lyman, J.: Performance Studies on the NOTS-UCLA Tracking Simulator: Effects of Selected Controller Configurations and Transfer of Training. Rep. no. 66-22, Dept. Eng., Univ. Calif., Los Angeles, 1966.
4. Smith, R. L.; Garfinkle, D. R.; Groth, H.; and Lyman, J.: Performance Studies on the NOTS-UCLA Tracking Simulator: Independent Effects of Error Magnification and Field of View. Rep. no. 66-23, Dept. Eng., Univ. Calif., Los Angeles, 1966.
5. Battig, W. F.; Nagel, E. H.; and Brogden, W. J.: The Effects of Error-Magnification and Marker Size on Bidimensional Compensatory Tracking. Amer. J. Psychol., vol. 68, 1955, pp. 585-594.
6. Bowen, J. H.; and Chernikoff, R.: The Relationship Between Magnification and Course Frequency in Compensatory Aided Tracking. Rep. no. 4913, U. S. Naval Research Lab., 1957.
7. Bowen, J. H.; and Chernikoff, R.: The Effects of Magnification and Average Course Velocity on Compensatory Tracking. Rep. no. 5186, U. S. Naval Research Lab., 1958.
8. Hartman, B. O.; and Fitts, P. M.: Relations of Stimulus and Response Amplitude to Tracking Performance. J. Exp. Psychol., vol. 49, 1955, pp. 82-92.
9. Seidenstein, S.; Chernikoff, R.; and Taylor, F. V.: The Relationship of Retinal-Gain Index to System Performance. Rep. no. 5548, U. S. Naval Research Lab., 1960.
10. Chernikoff, R.; Birmingham, H. P.; and Taylor, F. V.: A Comparison of Pursuit and Compensatory Tracking Under Conditions of Aiding and No Aiding. J. Exp. Psychol., vol. 49, 1955, pp. 55-59.
11. Chernikoff, R.; and Taylor, F. V.: Effects of Course Frequency and Aided Time Constant on Pursuit and Compensatory Tracking. J. Exp. Psychol., vol. 53, 1957, pp. 285-292.
12. Middleton, W. C.; and Howard, J. A.: Evaluation of Certain Control Systems With 40-mm Mount and TV Tracking. Report No. 59-25, Dept. Eng., Univ. Calif., Los Angeles, 1959.
13. Senders, J. W.: The Influence of Surround on Tracking Performance: I. Tracking on Combined Pursuit and Compensatory One-Dimensional Tasks With and Without a Structured Surround. Tech. Rep. no. 52-229, Part 1, WADC, 1953.

14. Lockard, R. B.: Evaluation Program of Tracking Controller Studies. Tech. note no. 304-7, Naval Ordnance Test Station, 1955.
15. Gottsdanker, R. M.: Prediction Motion With and Without Vision. Amer. J. Psychol., vol. 65, 1952, pp. 533-543.
16. Smith, R. L.; and Lyman, J.: The Tracking Loop: A Critical Review of Tracking and Related Sensorimotor Studies. Rep. no. 66-21, Dept. Eng., Univ. Calif., Los Angeles, 1966.
17. Fitts, P. M.; Marlowe, E.; and Noble, M. E.: The Interrelations of Task Variables in Continuous Pursuit Tasks: I. Visual-Display Scale, Arm-Control Scale, and Target Frequency in Pursuit Tracking. Res. bull. no. 53-54, Human Resources Research Center, 1953.
18. Noble, M. E.; Fitts, P. M.; and Marlowe, E.: The Interrelations of Task Variables in Continuous Pursuit Tasks: II. Visual-Display Scale, Arm-Control Scale, and Target Frequency in Compensatory Tracking. Res. bull. no. 53-55, Human Resources Research Center, 1953.
19. Andreas, B. G.; Green, R. F.; Smith, S.; and Spragg, S. D. S.: Two-Dimensional Compensatory Tracking Performance As a Function of Control-Display Movement Relationships, Positioning Vs. Velocity Relationships, and Miniature Vs. Large Stick Control. J. Psychol., vol. 48, 1959, pp. 237-246.
20. Birmingham, H. P.; and Taylor, F. V.: A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems. Rep. no. 4333, U. S. Naval Research Lab., 1954.
21. Gibbs, C. B.: The Continuous Regulation of Skilled Response by Kinaesthetic Feedback. Rep. no. 190-53, Medical Research Council, A. P. U., Cambridge, 1953.
22. North, J. D.; and Lomnicki, Z. A.: Further Experiments on Human Operators in Compensatory Tracking Tasks. Ergonomics, vol. 4, 1961, pp. 339-353.
23. Bahrick, H. P.; Bennett, W. F.; and Fitts, P. M.: Accuracy of Positioning Responses As a Function of Spring Loading in a Control. J. Exp. Psychol., vol. 49, 1955, pp. 437-444.

15. Force Feedback Compensation: A New Concept for Improved Manual Control System Performance

James H. Herzog and Richard W. Pew
University of Michigan

768 15916

In a previous study Notterman and Page (ref. 1) compared tracking performance between two equivalent dynamic systems. In one case the dynamics were represented in mechanical form on the control stick, but in the other case the dynamics were simulated on an analog computer. They found that system performance was uniformly better in the case of on-stick dynamics that the operator could feel. We believe that this principle may be generalized to the case in which the plant dynamics are not directly represented in mechanical terms by utilizing a force feedback compensation system that restores the correspondence between the mechanical feel of the control stick and the actual plant dynamics. In order to accomplish this, a mechanical analog of the differential equation describing the plant is implemented on the control stick. Then instead of using the control stick output position as the control variable applied to the plant, a signal proportional to the operator's output torque is employed. This effectively provides the appropriate mechanical feel characteristics without adding additional dynamic filtering in the forward loop.

This technique of force feedback compensation has been evaluated for the case of control of a lightly-damped second-order system. When the performance with the compensated system was compared with that of an isometric stick with analog plant dynamics, the compensated system was better by a factor of two to three. The difference was not reduced by extended practice with the two systems. It is believed that this concept has applicability in a number of areas of manual control. The technique seems especially suitable for complex plants where the benefit of supplementary sensory information can greatly improve performance.

REFERENCE

1. Notterman, J. M.; and Page, D. E.: Evaluation of Mathematically Equivalent Tracking Systems. *Perceptual and Motor Skills*, vol. 15, 1962, pp. 683-716.

16. Effect of Some Control System Nonlinearities on Single-Loop Compensatory Tracking*

Dunstan Graham
Systems Technology, Inc.

N68-15917

The last several years have seen a marked increase in detailed understanding of the effects on human-operator performance of changes in the task variables involved in single-loop compensatory tracking of random-appearing inputs. (See refs. 1 to 5.) This understanding has developed to the point where, with substantially linear manipulators and controlled element dynamics, the overall performance of the operator-machine system can be successfully modelled, and predictions of performance can be made.

At the same time there have been a number of ad hoc studies of compensatory tracking with one or more nonlinearities in the control loop (in addition to those possibly inherent in the operator). Any bases, however, for the generalization and extrapolation of the results of these studies are certainly not apparent (ref. 6). The main reason for the lack of well-defined, generalized knowledge of human-operator tracking capabilities in control systems with significant nonlinearities is related to the fundamental fact that in nonlinear systems the principle of superposition does not hold. Both the quantitative and even qualitative characteristics of systems which cannot be described as approximately linear are likely to be strong functions of inputs, specific conditions, characterizing parameters of the nonlinearities, and the location or ordering of the linear and nonlinear elements of the system. The large number of possible variables makes the experimental study of nonlinear feedback systems an extraordinarily onerous task.

Nevertheless, the investigation of manual control systems with significant nonlinearities continues to command research attention. Aircraft control systems, in particular, are characterized by a large number and variety of nonlinearities and it is important to delineate acceptable levels of typical aircraft control system nonlinearities (refs. 7 to 10). Figure 1 illustrates schematically some of the more significant nonlinearities which manifest themselves in the operation of a fully-powered aircraft longitudinal control system.

Incorporation of nonlinearities in a control loop may lead to at least three types of undesirable behavior atypical of a linear system (ref. 11):

- (1) Multiple equilibrium points (loss of static accuracy)
- (2) Stable limit cycle
- (3) Divergent instability for certain inputs or initial conditions

One might suppose that when nonlinearities are introduced into the system, the human operator would, if possible, adapt in such a way that the overall system does not exhibit these

*The research reported here was sponsored by the Behavioral Sciences Laboratory, Aerospace Medical Research Laboratories, under U. S. Air Force Contract AF 33(615)-1782.

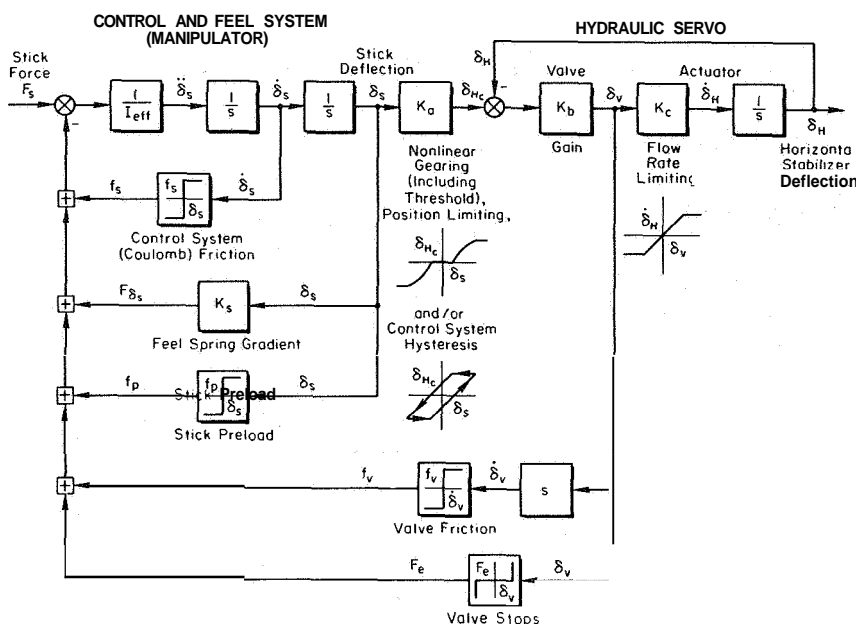


Figure 1.—Simplified aircraft longitudinal control system block diagram showing common nonlinearities.

"pathological" forms of behavior. The "linearization" which the operator applies or supplies can, in principle, take at least three forms:

- (1) Enclosing the nonlinearity in a high-gain feedback loop (ref. 12)
- (2) Inverting the nonlinearity in the operator's "effective controller"
- (3) Use of the general linearizing method, that is, dither (ref. 11)

Of course there might be situations in which these linearizations are impossible or ineffective and the performance of the system would then be expected to deteriorate. Data from an experimental series carried out to investigate the influence of nonlinearities on tracking performance can be adduced to illustrate both the linearization and the deterioration of performance when linearization is impossible.

EXPERIMENTAL APPARATUS AND FORCING FUNCTION

The principal item of equipment peculiar to this experimental series was the variable nonlinearity manipulator. It consisted of an aircraft-type seat and a light, center stick which could be moved side to side and fore and aft. Side-to-side motion was restrained only by a centering spring. Fore and aft motion of the stick was restrained by a variable gradient feel spring and otherwise by a miniature aircraft control system in which a variety of typical nonlinearities could be introduced and varied. Among these were control-system friction, preload, and hydraulic servo velocity limiting. The manipulator is described in more detail in reference 13.

Standard analog computer components were employed for the simulation of controlled element dynamics and for the accumulation of error scores, as well as other measures of performance, such as average absolute force and average absolute stick deflection.

An FM magnetic tape recorder was used to make a permanent record of selected signals so that power spectral densities and the cross spectra from which describing functions are derived could be computed on a digital computer.

The forcing function spectrum used throughout the experimental series was a so-called B-6 type (refs. 1 to 3) with a cutoff frequency of the main power components at 1.5 rad/sec and a root mean squared value of 0.5 inch as measured on the face of the operator's display.

A sum of 10 sinusoidal waves comprised the input forcing function. Five large amplitude waves were spaced approximately geometrically between 0.27 rad/sec and 1.5 rad/sec. Five additional waves, attenuated 20 dB, were spaced between 2.7 rad/sec and 20.0 rad/sec.

The difference between this input and the output of the controlled element, that is, the error, was displayed as the vertical displacement from center of a horizontal line segment on the face of a standard CRT. Gain of the controlled element had been set to a value which the operators agreed was "comfortable."

§SELECTED RESULTS

Some results of the experiments are presented in figures 2 to 5. Additional results with other nonlinearities and controlled elements are reported in reference 13.

Three measures of performance averaged over a 2-minute run are plotted in figures 2 and 3. First there is mean squared tracking error, normalized with respect to mean squared input $\bar{\epsilon}^2/\sigma_i^2$. Data points represent the mean value for each operator for four to eight runs at each condition. In the next lower part of each figure there is plotted the average absolute force, normalized with respect to average absolute error $|\bar{F}_s|/|\bar{\epsilon}|$. This is a measure of the operator's "force gain." Finally, the average absolute control deflection, normalized with respect to the average absolute error $|\bar{\delta}|/|\bar{\epsilon}|$, is also shown. This is a measure of the operator's "positioning gain."

It is clearly evident in figure 2 that, in these experiments, the introduction of comparatively large amounts of control stick friction did not significantly alter tracking performance. Mean squared error and average control deflection were unaltered by the addition of friction to the system; while, as one might perhaps expect, the average absolute force employed by the operator in performing the tracking task rose roughly in proportion to the increment in friction. Analogous results were obtained when 2 pounds of preload were combined with 2 pounds of friction and when 3 pounds of preload were combined with 4 pounds of friction as figure 2 shows. Similar results were also obtained as the feel spring gradient was varied from 0.6 lb/deg to 2.4 lb/deg with and without friction. Naturally, however, the average absolute force increased markedly with increases in the (linear) feel spring gradient.

These results were obtained with a controlled element comprising an integrator, that is, with what is sometimes called rate control. Other investigators, using either a pure gain controlled element or the simulated longitudinal dynamics of an airplane have reported a variety of results, such as

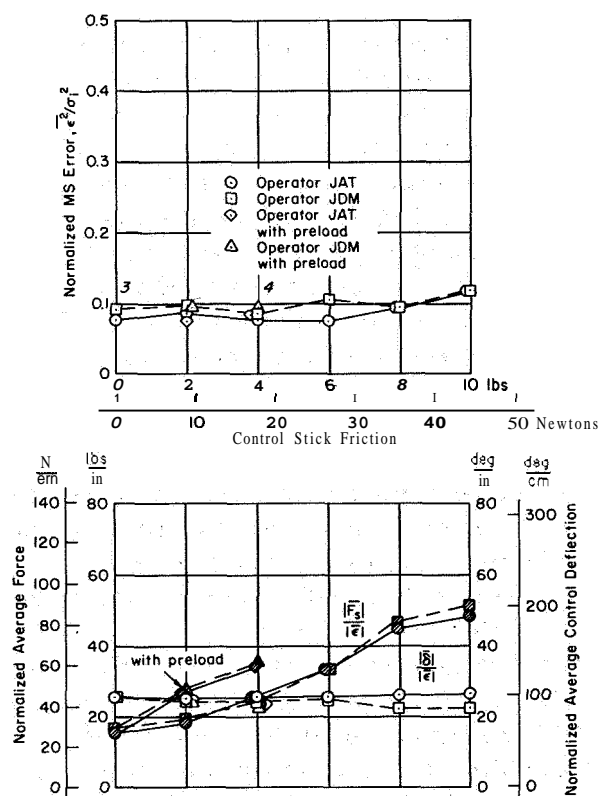


Figure 2.—Average performance with friction and preload; $Y_c = 0.2/s$; spring gradient = 0.6 lb/deg.

that friction made no significant difference, that low levels of friction improved tracking performance (in a moving cockpit or operator's station), that increased friction degraded tracking performance, or that differences existed but that a consistent trend was not discernible. (See ref. 6 for a summary of research on tracking with nonlinearities and ref. 14 for pertinent results not covered in ref. 6.) On the basis of the tests reported here, as well as tests in the same series made with simulated aircraft longitudinal dynamics and a very "difficult" controlled element comprising a first-order divergent instability, the evidence favors the conclusion that, in a fixed-base environment, neither friction nor preload has any significant effect on tracking performance. One suspects that in similar experiments, where a degradation in performance with increasing friction level was shown, it was because of a lack of motivation and/or training.

The data shown here, however, should not be interpreted to signify that high friction levels are in any sense "acceptable." Quite the opposite is true. Italic numerals adjacent to the mean squared error points on figures 2 and 3 correspond to the operator's opinion of the task. The scale runs from 1 \equiv "extremely pleasant" to 10 \equiv "absolutely uncontrollable." At the high friction levels represented in figure 2, the operators' comments were unprintable, and they did not correspond to any of the preassigned adjective descriptors.

Figure 2 presents strong circumstantial evidence for linearization by enclosing the nonlinearity in a high-gain feedback loop. Both the constant "positioning gain" $|\delta|/|\epsilon|$ and the increasing "force gain" $|\bar{F}_s|/|\epsilon|$ are what one would expect in a control loop adjusted to overcome the effect of the friction nonlinearity. There is further direct evidence of a wide band of dither frequencies between 4 rad/sec and 10 rad/sec in typical plots of the power spectral density of the operator's stick displacement output (ref. 13). One would not, however, expect that the operator could invert the friction nonlinearity in any effective way, and there is no indication that he does so. His characteristics as a high-gain positioning servo and dither generator are sufficient to overcome the friction nonlinearity.

Figure 3 illustrates the effect on tracking performance of a nonlinearity not amenable to linearization. This is control velocity limiting which typically arises because of flow saturation in hydraulic control surface actuators. It has been identified as the probable cause of pilot-induced oscillations in both longitudinal and lateral control (refs. 15 and 16). Pilot-induced oscillations are most often limit cycles of the nonlinear pilot-vehicle system.

The striking thing about figure 3 is the large difference in performance between the two operators. Operator JAT used much larger control forces, on the average, and achieved very significantly lower mean squared errors. In conducting the tests it was observed that operator JDM used a technique in which he simply did the best he could with the limited control velocity available. On the other hand, when operator JAT encountered the velocity limiting characteristic, he exerted himself to overpower the hydraulic servo. It is evident that he was successful with the model servo used in the experiments, although this tracking technique would not be useful in connection with the hydraulic actuators of actual aircraft.

In both cases, figure 3 shows a marked deterioration in performance as the control velocity limit, expressed in degrees per second of stick motion, was made lower. Note also the operator's rapidly deteriorating opinion of the task.

Comparable results were obtained with simulated aircraft longitudinal dynamics and the divergent instability as controlled elements (ref. 13). In these fixed-base experiments with an input, however, there was no evidence of self-sustained oscillations (limit cycles). This result corroborates the conclusions arrived at in reference 14. Control of the divergent instability with a time constant of 0.5 second was impossible for more than a few seconds at a time with the lowest value of control velocity limiting (15.1 deg/sec). These more "difficult" controlled elements further illustrated the practical impossibility of the operators adapting in an effective way to the velocity limiting nonlinearity.

It is interesting to compare the maximum acceleration capability in control of the integrator controlled element with control velocity limiting to the somewhat similar tests made by Bergeron, Kincaid, and Adams involving position limiting in control of a double integrator controlled element (ref. 17). At a value of control velocity limiting of 15.1 deg/sec, the results of figure 3 correspond to a maximum acceleration of $(15.1 \text{ deg/sec}) \times (0.2 \text{ in./sec/deg}) = 3.02 \text{ in./sec}^2$ which is shown in reference 17 to result in deteriorated performance. It would seem, however, that the effect of control velocity limiting is a more gradual reduction in capability than is the case with control position limiting.

The fact that control stick friction does not alter tracking performance, while control velocity limiting does, is also revealed in operator describing function measurements.

Figure 4 shows the average of two measurements of an operator's describing function for the case of tracking with an integrator controlled element and with 4 pounds of control stick friction. The data are actually plotted as the total open loop describing function $Y_p Y_c$ and are fitted with the "cross-over model" represented by the solid curves (ref. 3). Parameters of the cross-over model are in very close agreement with the definitive, linear results reported in reference 3, as well as with a linear, base test in the series of experiments described here. Similar results were obtained with the combination of friction and preload (ref. 13).

In Figure 5, on the other hand, the results are quite different. Here are plotted the describing function measurements for a single run while tracking with an integrator controlled element and with control velocity limiting at 15.1 deg/sec. Again, the data points represent the total open loop describing function $Y_p Y_c$ although the transfer function of the controlled element itself is shown as dashed lines so that the contribution of the operator and control system is made evident. Clearly there is here a reduction in the gain cross-over frequency ω_c corresponding to the increase in tracking error. More noticeable, however, is the phase dip with respect to the phase characteristic of a pure time delay. It appears at intermediate frequencies. This is not unlike the phase characteristic associated with the sinusoidal describing function for a velocity limited servo-motor (ref. 11). A direct comparison of the sinusoidal and random input describing functions is not, however, feasible. In any event the data points in figure 5 cannot reasonably be fitted with a cross-over model. Thus, the average performance in the frequency domain may be seen to be distinct from cases in which either the system is approximately linear or in which the operator is able to linearize the performance.

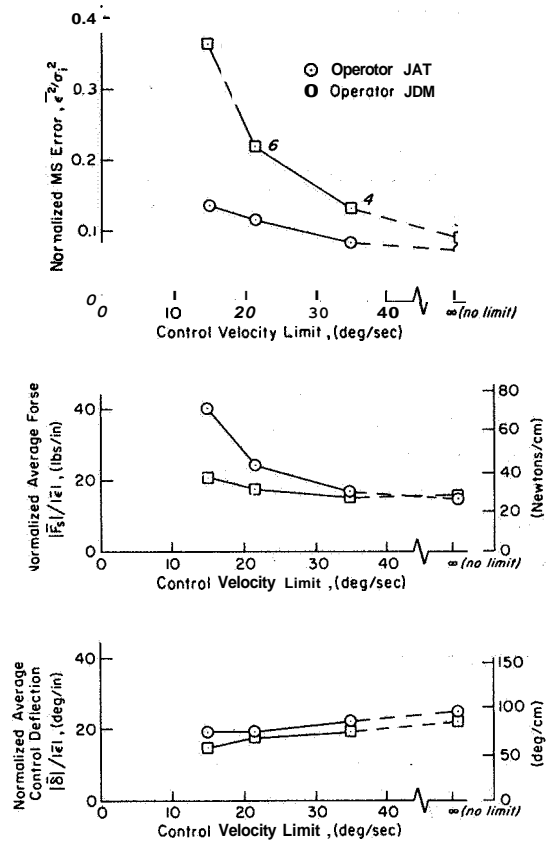


Figure 3.—Average performance with control velocity limiting; $Y_c = 0.2/s$.

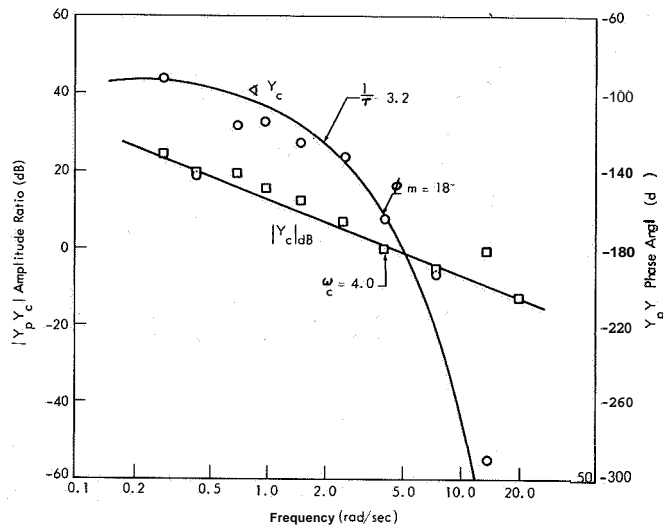


Figure 4.—Mean open-loop describing function; 4 pounds friction; $Y_c = 0.2/s$.

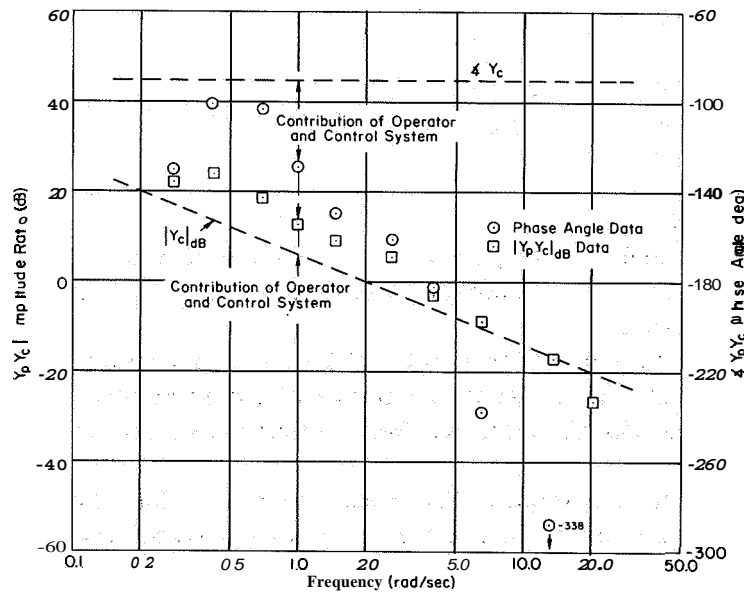


Figure 5.—Open loop describing function; $f 15.1 \text{ deg/sec}$ control velocity limiting; $Y_c = 0.2/s$.

CONCLUSIONS

On the basis of the data presented here, it can be concluded that the human operator in single-loop compensatory tracking of random appearing inputs will tend to linearize the performance of the system when this is possible. Both an adjustment as a high gain positioning servo and the use of dither are effective against friction and the combination of friction and preload. With the control velocity limiting nonlinearity, no linearization is possible and performance is degraded.

In the experiments described in this paper there was no evidence of self-sustained oscillations because of the nonlinearities. The same thing was true of other experiments in the series made with controlled elements thought to be very likely to produce this pathological behavior.

REFERENCES

1. Elkind, J. E.: Characteristics of Simple Manual Control Systems. TR-111, MIT Lincoln Lab., Apr. 6, 1956.
2. McRuer, D. T.; and Krendel, E. S.: Dynamic Response of Human Operators. WADC-TR-56-524, Oct. 1957.
3. McRuer, D. T.; Graham, D.; Krendel, E. S.; and Reisener, Jr., W.: Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15, July 1965.
4. Magdaleno, R. E.; and McRuer, D. T.: Effects of Manipulator Restraints on Human Operator Performance. AFFDL-TR-66-72, Dec. 1966.
5. McRuer, D. T.; and Magdaleno, R. E.: Human Pilot Dynamics With Various Manipulators. AFFDL-TR-66-138, Dec. 1966.
6. Wasicko, R. J.; and Magdaleno, R. E.: Effect of Nonlinearities on Human Operator Tracking Performance: A Review of the Literature. AMRL-TR-65-158, Oct. 1965.
7. McRuer, D. T.; Moser, J. E.; and Trudel, R. E. (eds.): The Artificial Feel System: BuAer rep. AE-61-4-V, Northrop Aircraft, Inc., May 1953.
8. Glenn, J. E.: Manual Flight Control System Functional Characteristics. IEEE Trans. on Human Factors in Electronics, vol. HFE-4, Sept. 1963.
9. Anon., Flying Qualities of Piloted Aircraft. MIL F-8785(ASG), Feb. 10, 1959.
10. Koven, W.; and Wasicko, R. J.: Flying Qualities Requirements for United States Navy and Air Force Aircraft. Presented at AGARD Specialists' Meeting (Brussels, Belgium), Apr. 10, 1961.
11. Graham, D.; and McRuer, D. T.: Analysis of Nonlinear Control Systems. John Wiley & Sons, Inc., 1961.
12. McRuer, D. T.; and Graham, D.: Pilot-Vehicle Control System Analysis. Guidance and Control-11, R. C. Langford, and C. J. Mundo, eds., vol. 13, Progress in Aeronautics and Astronautics, Academic Press, 1964.
13. Graham, D.: Research on the Effect of Nonlinearities on Tracking Performance. AMRL-TR-67-9, 1967.
14. Sadoff, M.: The Effects of Longitudinal Control-System Dynamics on Pilot Opinion and Response Characteristics as Determined from Flight Tests and from Ground Simulator Studies. NASA MEMO. 10-1-58A, 1958.
15. Ashkenas, I. L.; Jex, H. R.; and McRuer, D. T.: Pilot-Induced Oscillations: Their Cause and Analysis. Norair Report NOR-64-143, June 20, 1964.
16. Barnes, K. H.; and Robertson, F. C.: Study of the Effect of Increased Aileron Surface Rate Capability on Handling Qualities of the B-58A During Landing. Rep. FZE-4-051, General Dynamics/Fort Worth, Nov. 30, 1962.
17. Bergeron, H. P.; Kincaid, J. K.; and Adams, J. J.: Measured Human Transfer Functions in Simulated Single-Degree-of-Freedom Nonlinear Control Systems. NASA TN D-2569, 1965.

PRECEDING PAGE BLANK NOT FILMED

VI. DECISION PROCESSES II

189

17. An Asynchronous Pulse-Amplitude Pulse-Width Model

of the Human Operator N68-15918

M. J. Merritt and G. A. Bekey
University of Southern California

In order to develop a model for the behavior of a human operator performing a manual control task, it is necessary to make some assumptions concerning the operator's inputs and outputs. If it is assumed that the operator utilizes the input continuously and produces continuous outputs, there is a wide variety of techniques which can be used to develop complete human-operator models. These techniques include spectral analysis (refs. 1 and 2), multiple linear regression (ref. 3), and gradient search model identification (ref. 4). Another assumption which has been studied is that the human operator samples the inputs periodically and produces continuous outputs. The physiology of the optical, neuromuscular, and cerebral systems (ref. 5) supports this assumption. Sampling human operator models are difficult to identify due to the interactions between the sampling rate and time constants in the continuous portion of the model. For simple controlled elements (refs. 6 and 7) and step inputs (ref. 8) complete human operator models have been identified.

Although the physiology supports the use of sampled data models for human operators, many studies have produced no evidence of periodic sampling behavior. In a recent study (ref. 1) the power spectrum of the model remnant was examined for periodicities corresponding to sampling phenomena. No evidence of periodic sampling was found. However, experiments conducted at the University of Southern California have shown that small random perturbations about a nominal sampling interval tend to mask periodicities in the spectrum of the model remnant. The sampling behavior of human operators is certainly aperiodic and possibly controlled by a supervisory input monitor. This results in aperiodic input dependent sampling, which may also contain random variations in the sampling interval. If this is indeed the case, the model remnant would not contain strong periodicities.

When the dynamics of the controlled element contain two or more integrations, the performance of the human operator approaches that of a bang-bang system. In particular, the double integrator plant $1/s^2$ usually elicits pulse responses from human operators (refs. 9 to 11). A mathematical model to represent this output behavior could contain sampled inputs with continuous supervisory control of the sampling. This supervised sampling extends the periodic sampling of previous models (refs. 6 and 7) to aperiodic input-dependent sampling. The pulse nature of the output makes it possible to relate pulse events to decision surfaces in the error-phase space (refs. 5 and 12).

The object of this paper is to describe the development of a human-operator model which produces discrete outputs in response to continuously presented gaussian random inputs. Computer procedures for the complete identification of all model parameters are described.

STATEMENT OF THE PROBLEM

A block diagram of the compensatory tracking situation used in this study is shown in figure 1, and a portion of a typical tracking record is found in figure 2. An examination of the human-operator output (stick position) reveals a sequence of pulses which are roughly triangular in shape. For the purposes of this study, the actual human-operator output was converted to the idealized human-operator output, as seen in figure 3. The selection of symmetric triangular pulses as ideal human-operator pulses is arbitrary, and other pulse shapes can be used. Further it was decided to treat each pulse as a separate event, uncorrelated with previous pulses, in order to keep the structure of the pulse model as simple as possible. The use of preprogrammed pulse sequences (refs. 13 to 15) presents an opportunity for future extensions of the work.

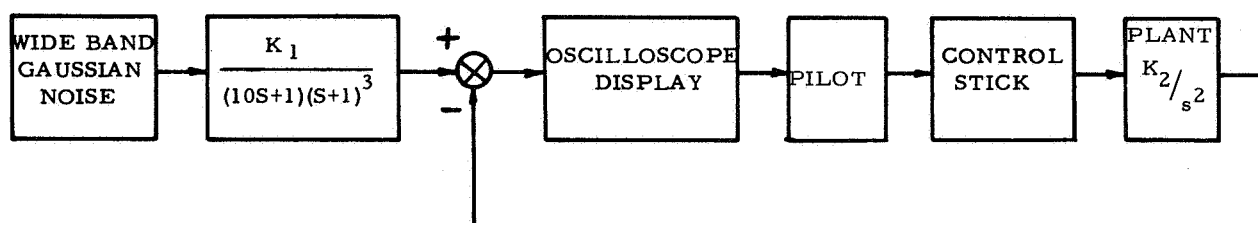


Figure 1.—Compensatory tracking system.

The idealized human-operator output can be represented by the following sequence: time of the pulse initiation, pulse amplitude, and pulse width. If a causal relationship exists between the transient human-operator inputs and the pulse outputs, the input record can be reduced to samples of the input in the vicinity of the pulse initiation. The objective of the present study is the determination of the relationships between these input samples and the pulse output.

Since each event is treated independently, short-term human-operator variations are easily computed. These variations are the difference between the model outputs and the actual human-operator outputs. The distribution functions of these variations can be obtained and, if desired, can be reinserted as model perturbations to produce a human-operator model statistically indistinguishable from the actual human operator. The distribution functions and their associated parameters (mean and moments) can be used as measures of performance and state of training.

THE EXPERIMENT

The compensatory tracking task shown in figure 1 was mechanized using a Beckman 2132 analog computer, an X-Y oscilloscope and side arm control stick. Operator distraction was minimized by placing the manual-control station inside a sound-proof enclosure with approximately 40 dB of audio attenuation. The operator wore an aircraft-type headset with a lip microphone for communication purposes. The operator sat in a chair without armrests facing the display oscilloscope. The control stick was adjustable in position and contained an integral arm rest. The operator adjusted the control stick and arm rest into a comfortable position. The oscilloscope was placed at eye level.

The double-integrator plant closely resembles an aircraft pitch axis. The input is elevator position and the output is altitude. In order to preserve this resemblance, the error display was a rotating needle corresponding to a glide-path indicator in an aircraft navigational/ILS display. A horizontal needle position represented zero error. The frequency

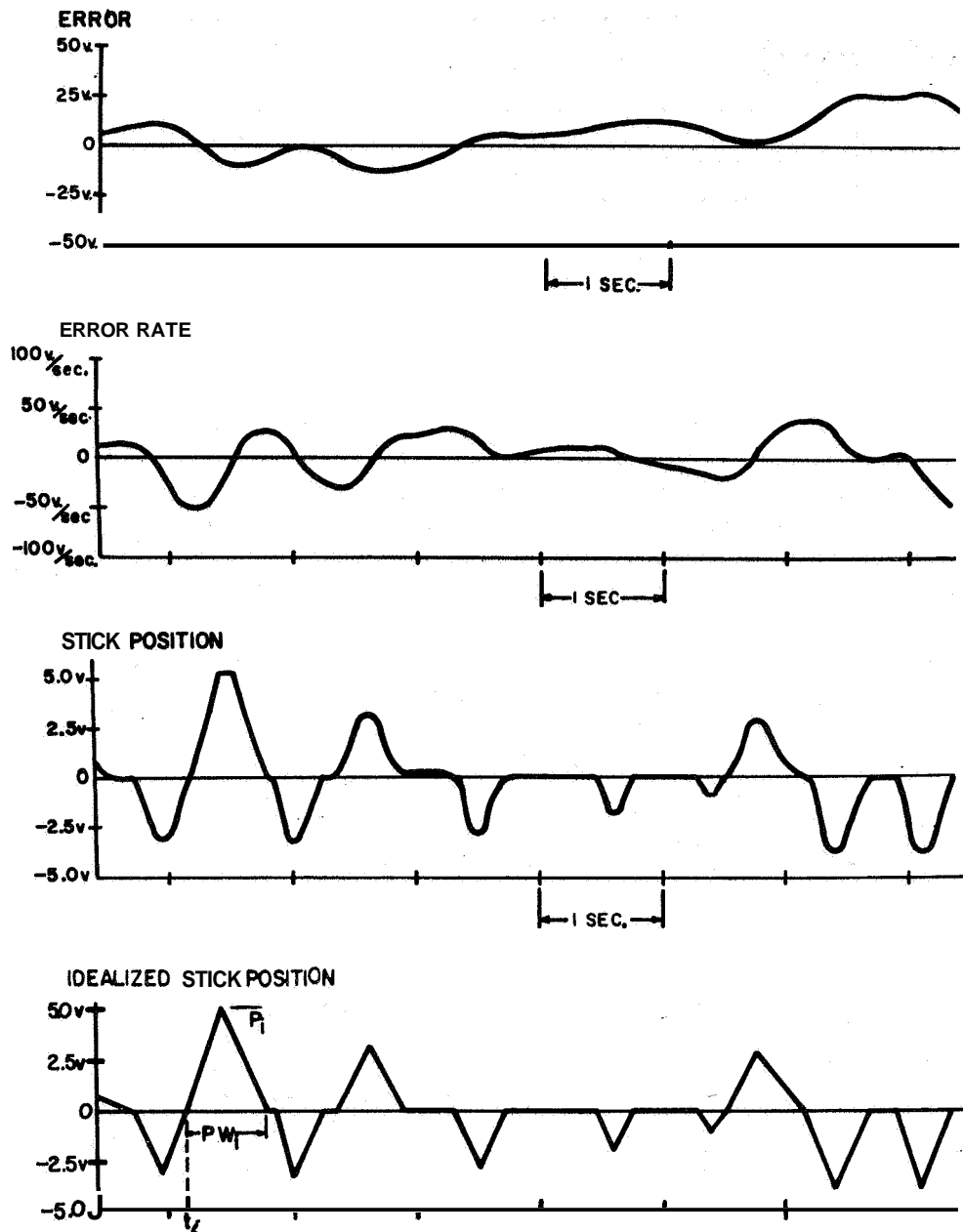


Figure 2.—Typical tracking record.

response problems associated with actual instruments were avoided by simulating the glide-slope needle with an oscilloscope containing a specially prepared edge-lighted reticle.

The control stick and oscilloscope were connected to the analog computer which converted the stick output to a voltage, computed the plant response, and generated the necessary X- and Y-axis signals for the error display. By solving some of the equations explicitly it was possible to obtain the error and its exact derivative. The inputs to the system were obtained by filtering the output of a low-frequency gaussian noise source. The filter transfer function was :

$$F(s) = \frac{K}{(10s+1)(s+1)^3} \quad (1)$$

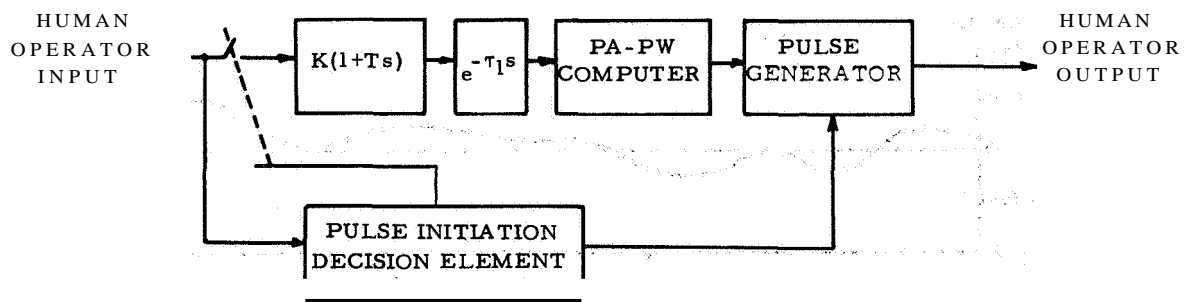


Figure 3.—General form of proposed human-operator model.

An FM magnetic tape recorder was used to record tracking data, which was later digitized and stored on a disk file for digital processing.

A single subject received approximately 20 hours of training over a period of 1 month. The training sessions consisted of 10 minutes of tracking with 10-minute rest periods. One of the last sessions was recorded on magnetic tape. From the 10-minute session approximately 3 minutes of data were subsequently digitized. The sampling interval utilized was 25 milliseconds or 40 samples per second.

The digitized data stored on the 1311 disk file were printed out and punched on IBM cards for permanent storage. The following data were punched on IBM cards:

- (1) The time of the pulse initiation
- (2) The time of the pulse termination
- (3) The peak amplitude of the pulse
- (4) The values of e and \dot{e} at the following times:
 - (a) One sample after the initiation of the pulse
 - (b) At the start of the pulse
 - (c) The five samples prior to the start of the pulse

HUMAN OPERATOR MODEL

In order to carry out modeling efforts, a basic model structure must be hypothesized. Intuitive concepts of human-operator behavior and an examination of the tracking record in figure 2 led to the model shown in figure 3. This model is based on the assumption that the operator generates an output consisting of a series of modulated pulses. The width and amplitude of these pulses are dependent on the error and error rate at some time prior to pulse initiation. Consequently, the next phase of the modeling effort consisted of investigating possible quantitative relationships between the input quantities (error and error rate) and the output pulse events.

PULSE AMPLITUDE RELATIONSHIPS.—The amount of pulse-amplitude modulation utilized by the human operator is evidenced in the distribution function of the pulse amplitudes shown in figure 4. Evidently the operator does not behave in a bang-bang mode. In fact, he utilizes a width range of amplitudes. Further work is needed to evaluate the significance of the asymmetry which appears in the data. Consider the idealized triangular pulse output record shown in figure 2. The relationship between the inputs and the pulse amplitude was postulated to be:

$$c_1 e(t_i - \tau_1) + c_2 \dot{e}(t_i - \tau_1) + c_3 = p_i \quad (2)$$

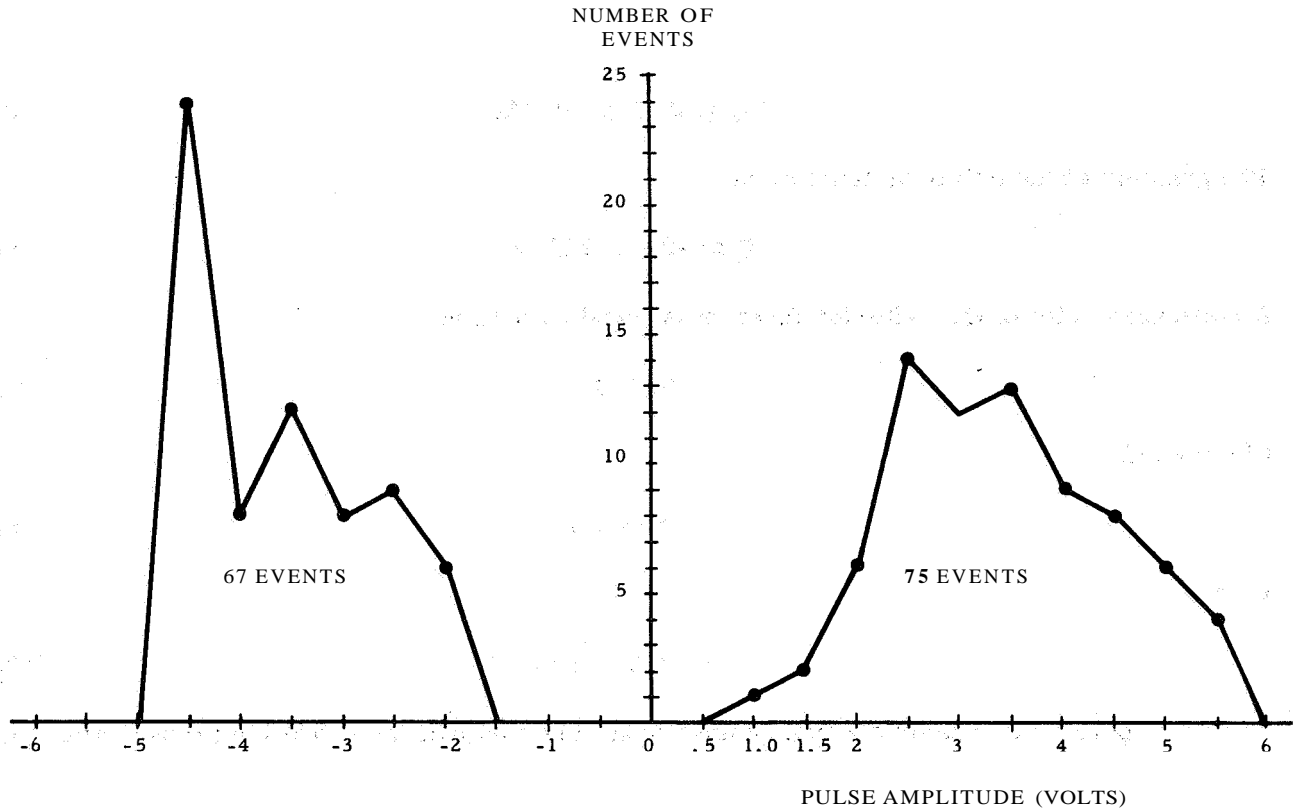


Figure 4.—Distribution of pulse amplitudes.

where c_1 , c_2 , and c_3 are constants, τ_1 is the time delay shown in figure 3, t_i are the times of pulse initiation as shown in figure 2, and e and \dot{e} are the human operator inputs.

The constants c_1 , c_2 , and c_3 are easily determined for fixed values of τ_1 . This procedure is natural since the digitizing of the tracking record with a sampling interval of $T=25$ milliseconds allows τ_1 to take on only discrete values which are multiples of T . Let τ_1^* be one of these fixed values of τ_1 . If there are N events (human-operator pulses) then there are N equations in three unknowns:

$$\left. \begin{aligned} p_1 &= c_1 e(t_1 - \tau_1^*) + c_2 \dot{e}(t_1 - \tau_1^*) + c_3 \\ p_2 &= c_1 e(t_2 - \tau_1^*) + c_2 \dot{e}(t_2 - \tau_1^*) + c_3 \\ &\vdots \\ p_N &= c_1 e(t_N - \tau_1^*) + c_2 \dot{e}(t_N - \tau_1^*) + c_3 \end{aligned} \right\} \quad (3)$$

or, in vector form

$$p = Ac \quad (4)$$

where A is an $N \times 3$ matrix of the values of e and \dot{e} . A least-squares error criterion is:

$$\phi = (p - Ac)'(p - Ac) \quad (5)$$

which is a positive definite quadratic function of the parameter vector, c . In expanded form,

$$\phi = p'p - 2c'A'p + c'A'Ac \quad (6)$$

The gradient of the criterion function is

$$\nabla_c \phi = -2A'p + 2A'Ac \quad (7)$$

A stationary point of the criterion function is found by setting

$$\nabla_c \phi = 0 \quad (8)$$

which yields

$$A'Ac = A'p \quad (9)$$

Finally

$$c = [A'A]^{-1} A'p \quad (10)$$

The parameter vector c which results from this computation represents those values of c_1 , c_2 , and c_3 which produce the best least-squares fit between the pulse amplitude predicted by the model

$$PM_i = c_1 e^{(t_i - \tau_1^*)} + c_2 e^{-(t_i - \tau_1^*)} + c_3 \quad (11)$$

and the pulse amplitudes p_i for a given value of τ_1 , for instance, τ_1^* and $i=1, 2, \dots, N$.

The criterion function ϕ is a measure of the correlation between the model relationship and the experimental data. In order to gain more insight into the problem, the values of PM_i may be plotted against the actual pulse amplitude p_i for each of the N pulse events. If the human operator were invariant with time and the model were an exact representation, the plotted points would lie on a straight line with a slope of 1, corresponding to a criterion function of zero. There are two ways in which the optimal value for τ_1 can be selected from the set of values τ_1^* used in the linear regression routine above. The first is to examine the point plots just described for various values of τ_1^* and to select the plot which produced the best visual approximation to a perfect straight line. The second method is to select that value of τ_1^* which produces the minimum criterion function ϕ . Both methods will be used below in analyzing and interpreting the results from the least-squares linear-regression analysis.

COMPUTER RESULTS.—A least-squares linear-regression routine mechanizing the procedure described above was written. In order to make possible the investigation of asymmetry in the human operator's response, the positive and negative pulses of the tracking record were analyzed separately. The results of these computations are tabulated in table 1 and plotted in figures 5 and 6.

A plot of the criterion function ϕ versus τ_1^* is found in figure 6. From this figure it can be seen that the minimums are well defined. The scatter plot of model-pulse amplitude PM_i versus the actual pulse amplitude p_i is found in figure 5. The groupings are quite close to the ideal unity slope line. The optimum values for positive and negative pulses, respectively are $c_1 = -0.062$ and -0.066 ; $c_2 = -0.125$ and -0.143 ; $c_3 = 2.16$ and -1.99 ; and $\tau_1 = 0.050$ and 0.100 . The symmetry observed in these values contrasts with the large differences in optimum

TABLE 1.—POSITIVE AND NEGATIVE PULSES OF TRACKING RECORD

| τ_1^* | c_1 | | c_2 | | c_3 | | $\phi \times 10^{-6}$ | |
|------------|--------|--------|--------|--------|-------|-------|-----------------------|-------|
| Polarity | + | - | + | - | + | - | + | - |
| 0.125 | -0.060 | -0.065 | -0.139 | -0.145 | 2.16 | -2.02 | 0.507 | 0.220 |
| 0.100 | -0.061 | -0.066 | -0.134 | -0.143 | 2.16 | -1.99 | 0.496 | 0.217 |
| 0.075 | -0.062 | -0.065 | -0.130 | -0.137 | 2.16 | -2.00 | 0.495 | 0.222 |
| 0.050 | -0.062 | -0.064 | -0.125 | -0.131 | 2.16 | -2.04 | 0.492 | 0.228 |
| 0.025 | -0.059 | -0.061 | -0.118 | -0.123 | 2.19 | -2.10 | 0.498 | 0.242 |
| 0 | -0.057 | -0.060 | -0.112 | -0.116 | 2.22 | -2.14 | 0.504 | 0.253 |
| -0.025 | -0.057 | -0.060 | -0.106 | -0.108 | 2.25 | -2.20 | 0.509 | 0.263 |

criterion function, $\phi_+ = 0.492$ and $\phi_- = 0.217$. In other words, the positive pulses produced poorer correlation than the negative pulses. This may be a result of one or more of the following factors: arm motion asymmetry associated with the side arm control stick, incomplete training, or the tendency of the operator to prefer certain portions of the error phase plane (also a well-known result of incomplete training).

An analysis of the differences between human operator amplitude and the corresponding model pulse amplitude serves two purposes. The accuracy of the model may be measured, and the distribution function of the human operator variations may be determined. The mean and standard deviation of these functions are presented in table 2.

The amplitudes of human operator pulses range from 0.5 to 6.0 volts, with a distribution function as shown in figure 4. One standard deviation represents about a 40-percent error. Alternatively, noise of the same mean and standard deviation or the same distribution function could be added to the pulse-amplitude model output.

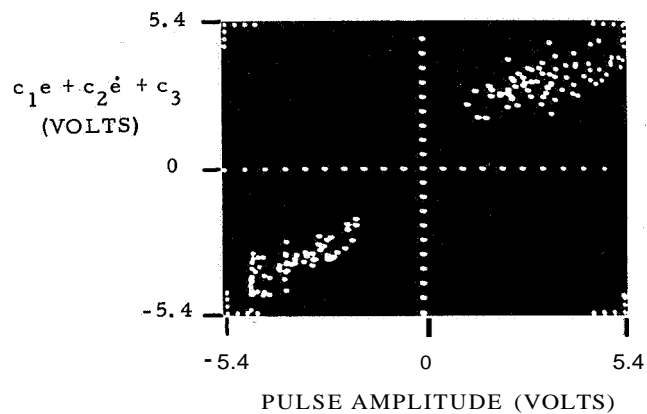
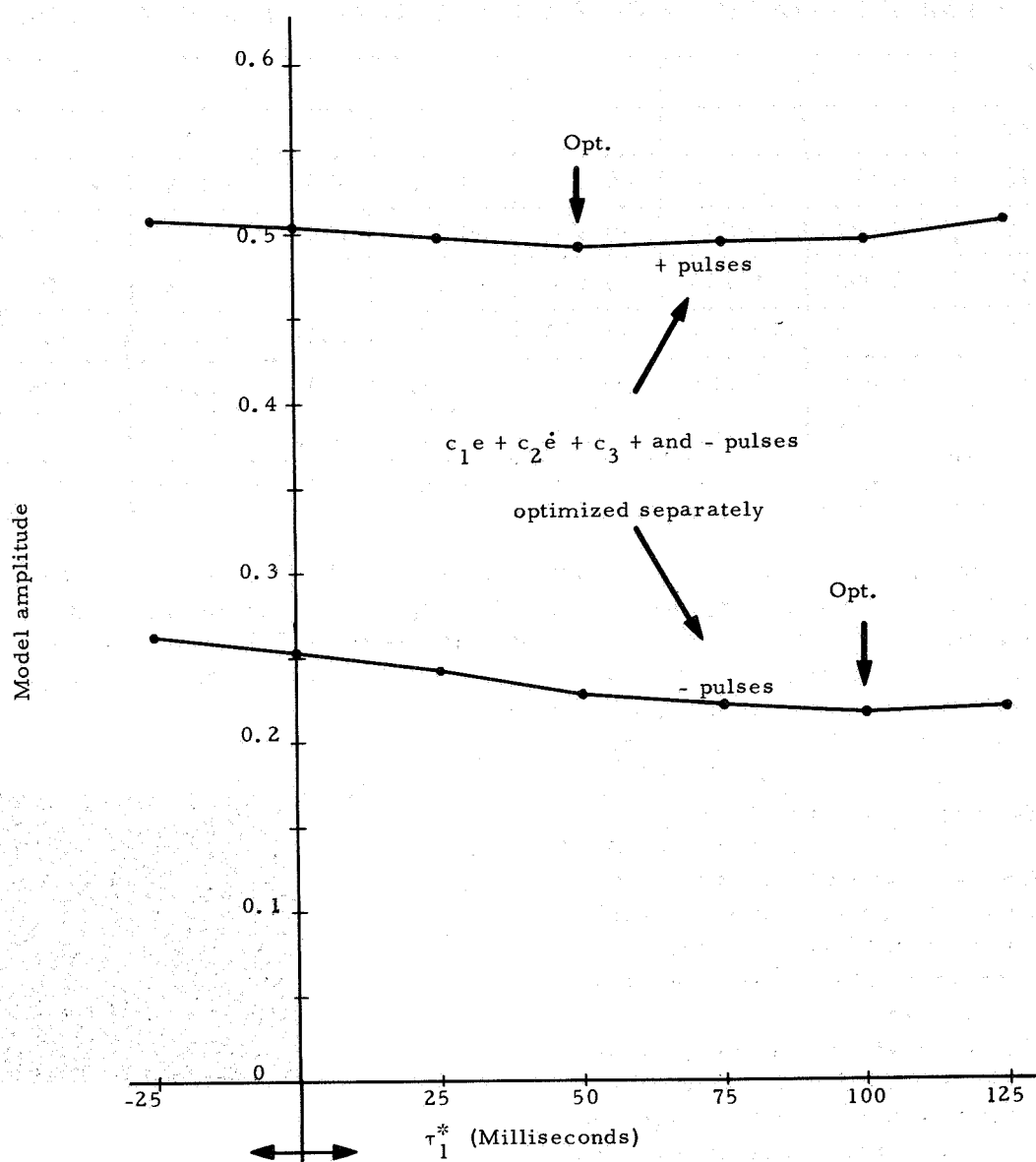


Figure 5.— $c_1e + c_2e + c_3$ versus pulse amplitude with positive and negative pulses optimized separately.

PULSE WIDTH RELATIONSHIPS.—Preliminary analysis of the tracking records led to the hypothesis that the pulse width was proportional to the pulse amplitude. The results presented subsequently indicate that this is not necessarily the best model structure. A future study might consider the pulse width as another degree of freedom which is determined independently of pulse amplitude. A plot of pulse width versus pulse amplitude for each of the 150 events is found in figure 7. The pulse width appears to be independent of pulse amplitude. Least-squares linear regression was used to determine the optimal parameter values in the following equation:

$$c_1p_i + c_2 = pw_i \quad (12)$$



| | + | - |
|-------------------------------------|------|------|
| Mean, volts. | 0 | 0 |
| Standard deviation, volts | 0.80 | 0.58 |

human operator and the model outputs were computed. The mean and standard deviation of these differences are presented in table 3.

The values of pw_i range from 0.3 to 0.9 second. One standard deviation represents a maximum error of 16 percent. As with the pulse amplitude model, the output of the pulse width model could be perturbed by random noise possessing the same distribution functions.

PULSE INITIATION.—The results of the pulse-amplitude model clearly demonstrate the ability of the human operator to estimate the derivative of displayed signals. The pulse-initiation decision model is based on this ability and the intuitive feeling that the human operator utilized a control policy which results in a relatively simple decision surface in the error-phase space. Inspection of the error-phase-plane trajectories immediately preceding pulse initiation led to the observation that the human operator utilizes the favorable error rate in the second and fourth quadrants by allowing the system to coast until the magnitude of the error is sufficiently small. If at that time the error rate is still large, a new pulse event occurs. A further observation is that an error/error-rate dead zone exists inside of which no pulse events are generated. This is consistent with other human-operator tracking experiments (refs. 6 and 12).

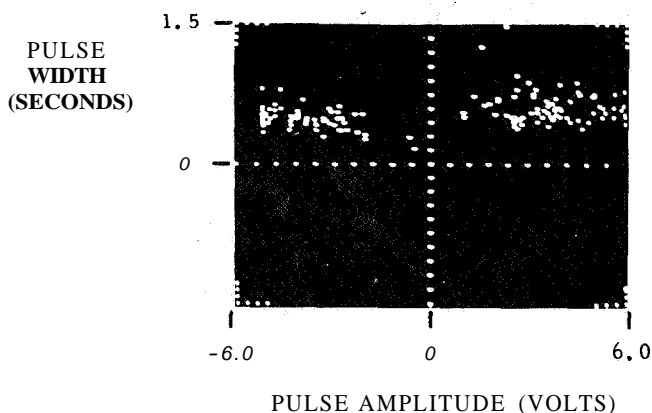


Figure 7.—Pulse width versus pulse amplitude.

TABLE 3.—MEAN AND STANDARD DEVIATION OF DIFFERENCES BETWEEN HUMAN OPERATOR AND MODEL OUTPUTS

| | Pulse polarity | |
|-----------------------------------|----------------|------|
| | + | - |
| Mean, sec | 0 | 0 |
| Standard deviation, sec | 0.05 | 0.04 |

Based on the above observations, it was hypothesized that the initiation of an output pulse by the human operator is a complex process which consists of several phases as follows :

- (1) At some time near the completion of an output pulse, monitoring of e and \dot{e} by a decision element starts.
- (2) When the error trajectory enters preselected regions of the phase plane, a decision to produce a pulse is made.
- (3) Some time later a pulse is generated.

The treatment of each pulse as an individual event places restrictions on the decision element. A number of investigators (refs. 5 to 7, 12, and 14) have determined human-operator reaction times to be between 150 and 300 milliseconds. In figure 8 is found the distribution function of the times between pulses. There are a large number of pulses spaced less than 200 milliseconds apart. These pulses probably correspond to preprogrammed pulse sequences (refs. 13 and 15). The human operator generates single pulses (rate corrections)

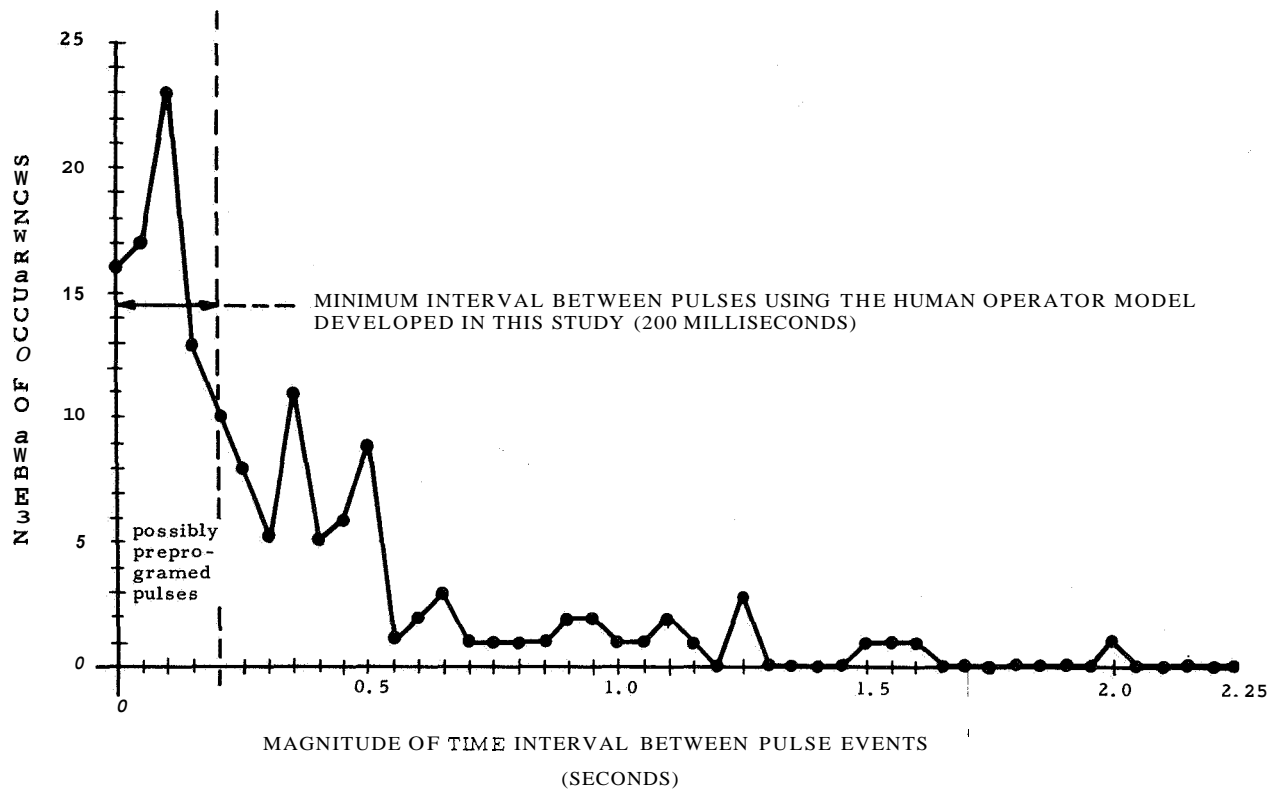


Figure 8.—Distribution function of magnitude of time interval between pulse events (from the termination of the i th event to the start of the $i + 1$ st event).

and pulse sequences (position corrections). A continuation of this study would include a pulse program decision element as described by Bekey and Angel (ref. 13). With this limitation in mind, the identification of the decision element was undertaken.

Inspection of the error phase plane trajectories led to the pulse initiation model shown below in figure 9.

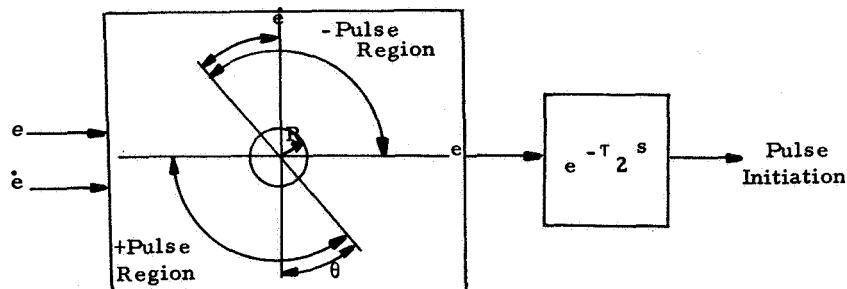


Figure 9.—Pulse initiation decision element.

A criterion function selected in accordance with the verbal description above is

$$\phi(R, \theta) = \frac{1}{150} \sum_{i=1}^{150} \left[\tau_2 - (t_i - t_{im}) \right]^2 \quad (13)$$

where:

- R the radius of a circular threshold region in the error phase plane (see fig. 9)
 θ an angle defined in figure 9
 t_i the time of initiation of the i th pulse
 t_{im} the time at which the region of the error phase plane defined by R and θ in figure 9 is entered (i.e., a decision to generate the i th pulse is made by the model)
 τ_2 the time delay between decision and initiation as defined by

$$\tau_2 = \frac{1}{150} \sum_{i=1}^{150} (t_i - t_{im}) \quad (14)$$

The criterion function, equation (13), is the variance of the delay time.

A systematic study of the R, θ parameter plane was conducted. The optimum parameter values were found to be $R=4.0$ volts, $\theta=35^\circ$, and $\tau_2=0.200$ second. The distribution of differences between model pulse initiation and human-operator pulse initiation times is found in figure 10. The peak at 100 milliseconds late is probably due to preprogrammed pulse sequences. The use of more complex decision elements is clearly indicated. However, these results demonstrate the applicability of discrete decision elements to the development of input-dependent-sampling human-operator models.

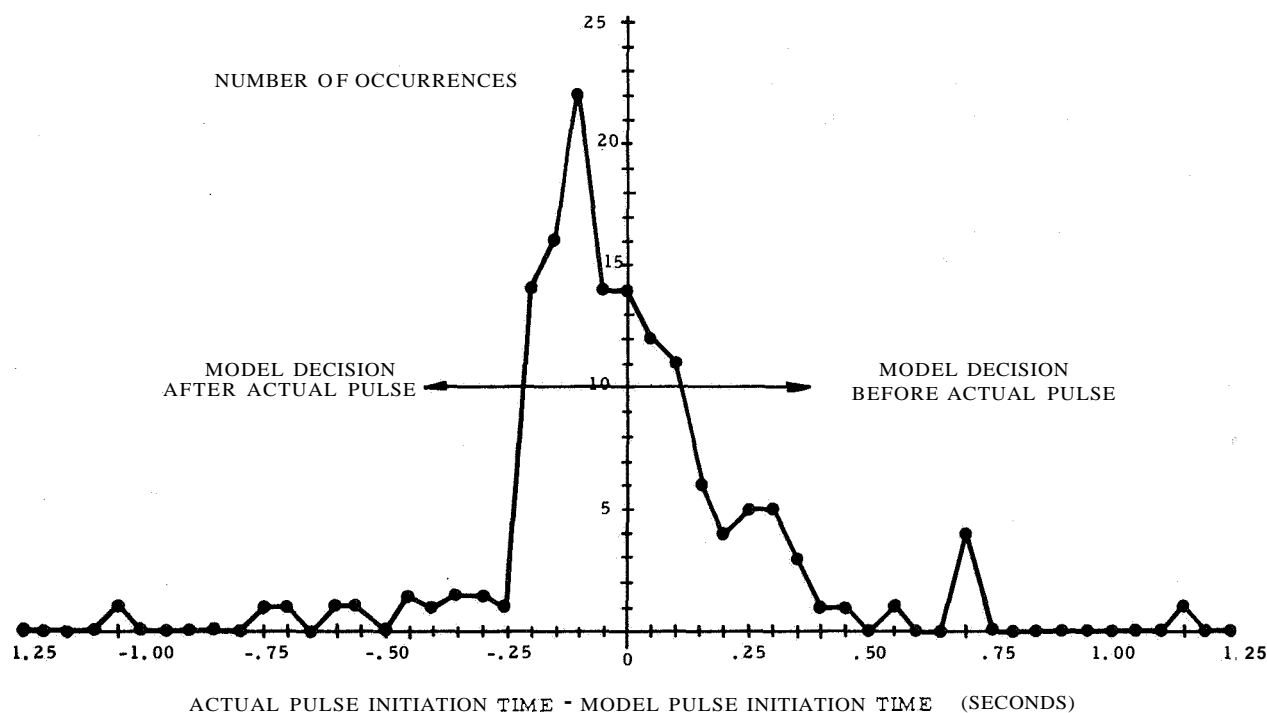


Figure 10.—Distribution function of pulse initiation model error.

COMPLETE HUMAN OPERATOR MODEL

The complete human-operator model is shown in figure 11. The parameters of this model have been identified as described above. The completely identified human-operator model for one well-trained subject is found in figure 12.

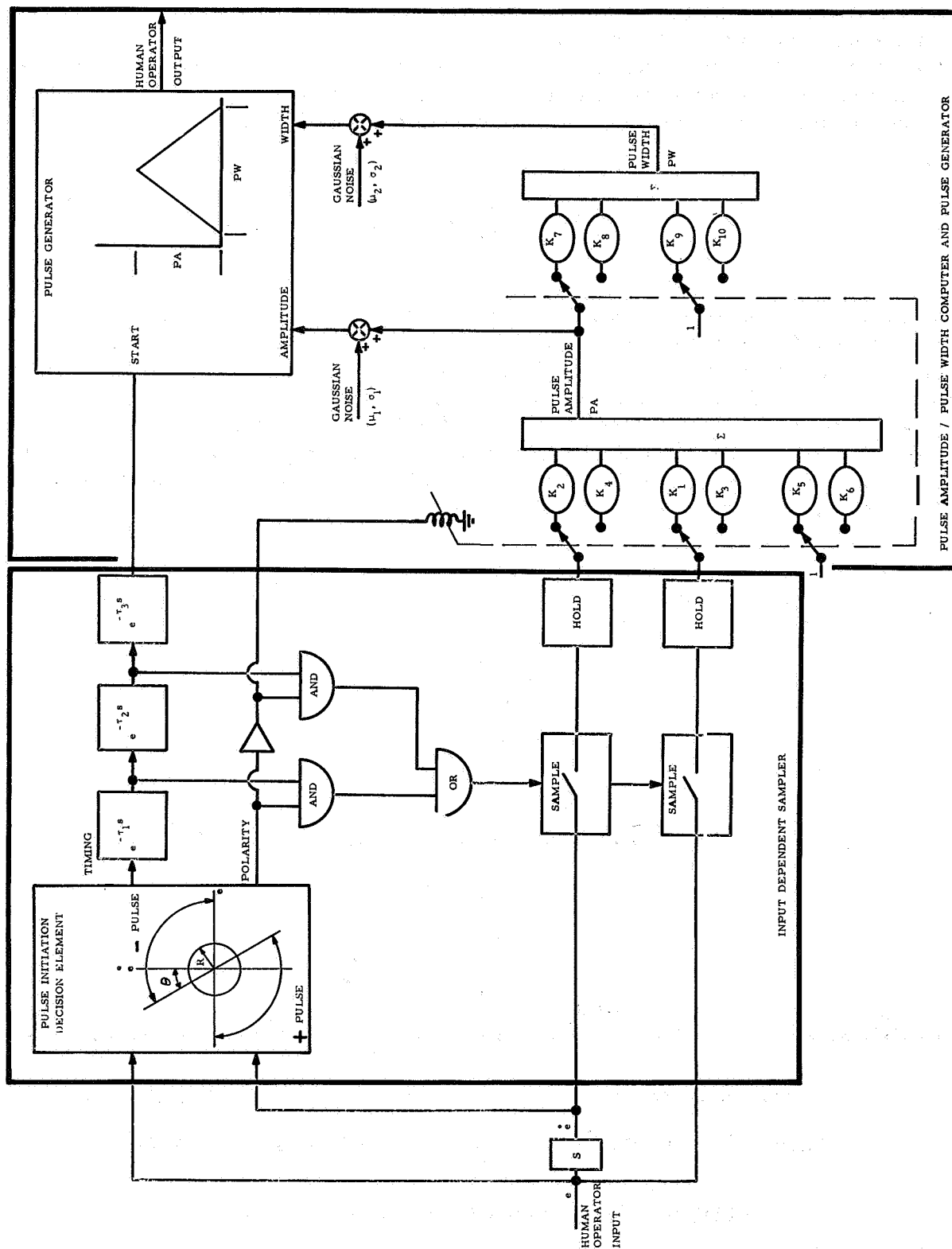


Figure 11.—Complete human-operator model.

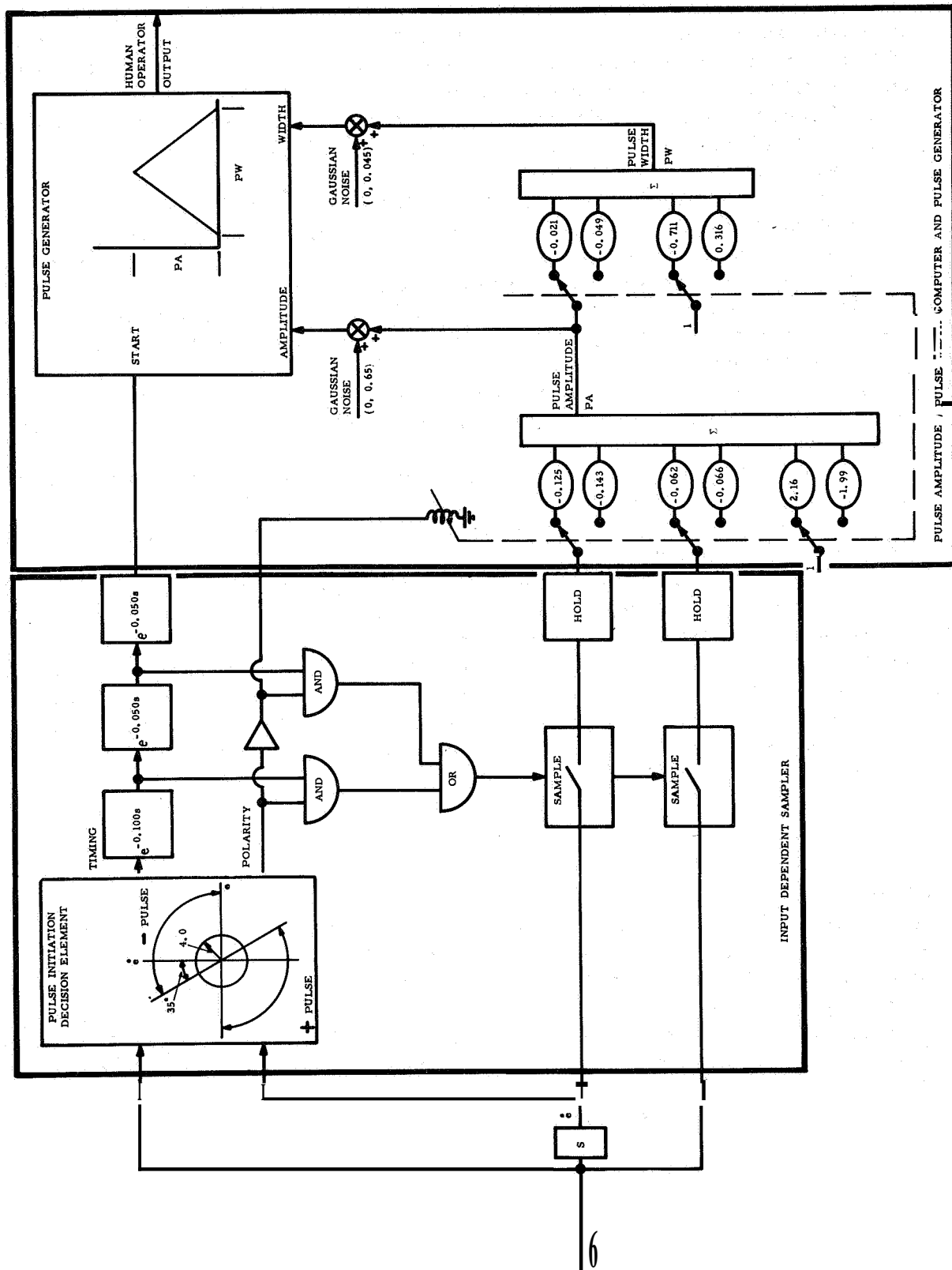


Figure 12.—Completely identified human-operator model for one well-trained subject.

SUMMARY OF RESULTS AND CONCLUSIONS

The parameters of the human-operator model shown in figure 11 were identified from experimental data taken from one subject in an advanced state of training. No records were made of the error or measures of the error as a function of training. The computational results brought to light a number of interesting results:

(1) The delay time τ_2 between the model's decision to pulse and the actual event was, 200 milliseconds. The value is within the range of reaction times reported in the literature (refs. 1, 5 to 7, 12, and 14).

(2) The numerical values for the time delays in figure 12 lead to the following sequence: (1) A decision is made to generate a pulse, followed by (2) a pause of 100 to 150 milliseconds, (3) $e(t)$ and $\dot{e}(t)$; we sampled (4) during the next 50 to 100 milliseconds; the amplitude and width of the pulse are computed; (5) the pulse is generated.

(3) The pulse-amplitude and pulse-width models for negative pulses produce better correlations with the experimental data than the models for positive pulses. This is clearly apparent in the scatter plot, figure 5, and in the values of the criterion function ϕ , in figure 6. This may be the result of incomplete training, the design of the side-arm controller used, the position of the subject's arm relative to the controller, or a characteristic of the particular human operator in this experiment.

(4) The pulse amplitude models for positive and negative pulses are quite similar, despite considerable asymmetry in pulse amplitude distributions (fig. 4).

(5) The results presented in figure 8 strongly indicate that human operators utilize some preprogramed pulse sequences.

(6) If the differences between model results and experimental tracking data are viewed as the result of short term human operator variations, then the statistics of the human operator variations are easily determined (tables 2 and 3).

From the present study it is not feasible to determine whether the model errors observed are random or functionally dependent on the human-operator inputs and input-output history. Further studies should include preprogramed pulse elements, more complex error-phase-plane decision surfaces, and the effects of training on the model parameters and their associated distribution functions.

REFERENCES

1. McRuer, D. T.; Graham, D.; Krendel, E. S.; and Reisner, W., Jr.: Human Pilot Dynamics in Compensatory Systems. Tech. rep. no. AFFDL-TR-65-15, July 1965.
2. Todosiev, E. P.; Rose, R. E.; Bekey, G. A.; and Williams, H. L.: Human Tracking Performance in Uncoupled and Coupled Two Axis Systems. Rep. 4380-6003-R0000, TRW Systems, 1965.
3. Wierwille, W. W.: A Theory for Optimal Deterministic Characterization of Time Varying Human Operator Dynamics. Cybernetics, Human Factors, IEEE Convention Record, part 6, 1965.
4. Meissinger, H. F.; and Bekey, G. A.: An Analysis of Continuous Parameter Identification Methods. Simulation, Feb. 1966.
5. Pew, Richard W.: Temporal Organization in Skilled Performance. Tech. Rep. 02814-11-T, Univ. Michigan, 1963.
6. Bekey, G. A.: Sampled Data Models of the Human Operator in a Control System. Ph. D. dissertation, University of California, Los Angeles, 1962.
7. Bekey, G. A.: The Human Operator as a Sampled Data System. IRE Trans., vol. HFE-3, no. 2, Sept. 1962.

8. Fu, K. S.; and Knoop, D. E.: An Adaptive Model of the Human Operator in a Control System. Rep. TREE 64-15, Purdue University, 1964.
9. Elkind, J. I.; Kelley, J. A.; and Payne, R. A.: Adaptive Characteristics of the Human Controller in Systems Having Complex Dynamics. Proceedings of the Fifth National Symposium on Human Factors in Electronics, 1964.
10. Kelley, Charles R.: Design Applications of Adaptive (Self-Adjusting) Simulators. Second Annual NASA-University Conference on Manual Control, NASA SP-128, 1966.
11. Young, L. R.; and Meiry, J. L.: Manual Control of an Unstable System With Visual and Motion Cues. IEEE Convention Record, part 6, 1965.
12. Young, L. R.; and Stark, L.: Biological Control Systems - A Critical Review and Evaluation. NASA CR-190, 1965.
13. Bekey, G. A.; and Angel, E. S.: Asynchronous Finite State Models of Manual Control Systems. Elect. Sci. Lab. Rep. USCEE-160, Univ. of Southern California, 1966.
14. Pew, Richard W.: Performance of Human Operators in a Three-State Relay Control System with Velocity-Augmented Displays. IEEE Trans. on Human Factors in Electronics, vol. 7, no. 2, June 1966.
15. Tomovic, R.; and McGhee, R. B.: A Finite State Approach to the Synthesis of Bioengineering Systems. IEEE Trans. on Human Factors in Engineering, vol. 7, no. 2, June 1966.

18. A Quantal Model for Human Tracking

E. C. Poulton
Applied Psychology Research Unit,
Cambridge, England

N68

15919

I am groping for a computer program which actually mimics the human in tracking inputs whose frequencies extend up to at least 30 cycles per minute. Suppose a person were handed 10 pairs of paper records, one of each pair produced by a man and the other produced from the same input by the program, and were asked to sort the members of each pair into one of **two** piles, according to whether it was produced by the man or by the program (ref. 1). With a program which mimics the man really well, it should be impossible to do this better than by chance.

This is a fairly strict criterion. A program could produce a similar output spectrum to the man, and also match his rms error, and yet look quite different from the man's output. While there may be human-operator models which can pass this test with high-frequency inputs, I have yet to discover them.

Figure 1 is an original figure from Tustin's 1947 paper (ref. 2). It represents **two** attempts by the same man to track the same piece of forcing-function input shown at the top of the figure. In my hypothetical test, the output of the program would have to match these two records as well as they match each other. Perhaps more difficult, the output of the program would have to contain no idiosyncrasy which would label it at once as nonhuman.

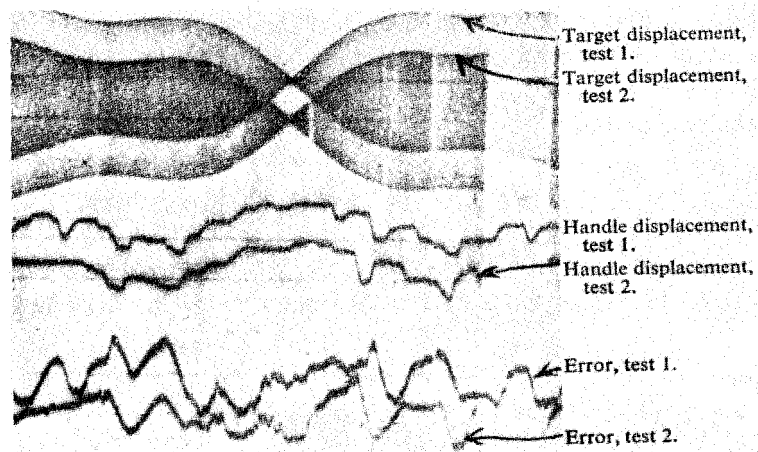


Figure 1.—Attempts to track a piece of forcing-function information (from ref. 2).

DISCRETE DECISIONS

Probably no program can mimic the human really well with high-frequency inputs unless it embodies discrete decisions **two** to three times per second. Craik (refs. 3 and 4) hit on this point by examining tracking records like those in figure 1, which are probably from an unpracticed man. With practice the signs of intermittency tend to disappear from the tracking records of low-frequency inputs like figure 1.

To explain what I mean by discrete decisions, consider someone tapping as quickly as possible. He may be able to tap for short periods at a rate of 7 to 9 taps per second. He can

achieve this fast rate because his tapping is preprogrammed. Each tap is not preceded by a decision to tap. He simply makes one decision to start tapping, decisions perhaps two to three times per second to continue tapping, and a final decision to stop tapping. If one suddenly tells him to stop or suddenly pulls away the piece of paper his pencil is tapping on, he will make one or two more taps before he can stop. Behavior of this kind has been reported by Noble, Fitts, and Warren (ref. 5) and by Poulton (ref. 6) in tracking simple sine waves with frequencies of 60 cycles **per** minute and above. Similar preprogrammed behavior has recently been reported by Pew (ref. 7) in two-state tracking with an acceleration control system.

When each response has to be preceded by a decision, a man cannot achieve response rates above 2 to 3 responses per second. To insure that a decision precedes each response, the required response has to be changed upon a proportion of occasions. Leonard and Newman (ref. 8) required a single response on 68 percent of occasions, and four other responses each on 8 percent of occasions in random order. After 2 weeks of practice at this task, groups of young naval enlisted men averaged only 2.5 responses per second. I believe that behavior of this kind is analogous to tracking an irregular input containing high frequencies.

In order to present the model, I will assume that the human tracker makes discrete decisions every half a second. The aim of each decision is to bring the error in position and rate to zero half a second later. The exact value of the time interval between decisions, like the exact values of the various weighting functions introduced subsequently, has to be determined by trial and error while using the program to match a particular piece of tracking output. Here the experimenter is faced with much the same problem as the processman in a chemical or petrochemical plant, who has to adjust his parameters to optimize the plant output. In this paper I shall be concerned only with the formal nature of the model, not with exact numerical values.

PURSUIT AND COMPENSATORY TRACKING

The task which faces the man in tracking a high-frequency forcing function with a pursuit display is quite different from the task which he faces with a compensatory display

(fig. 2). It is also much easier. Models which mimic the man's behavior really well must therefore be different for the two displays.

The pursuit display shows separately the forcing function input and the man's control system output. The man can see the present position and rate of the input. And as he watches the input, he can learn its statistical properties, such as its mean frequency and amplitude. The man can also see the position and rate of his control system output. Thus, he can have a visual model of his task in terms of movements in the display.

In contrast, the compensatory display shows the man only his error, the difference between the forcing-function input and his control

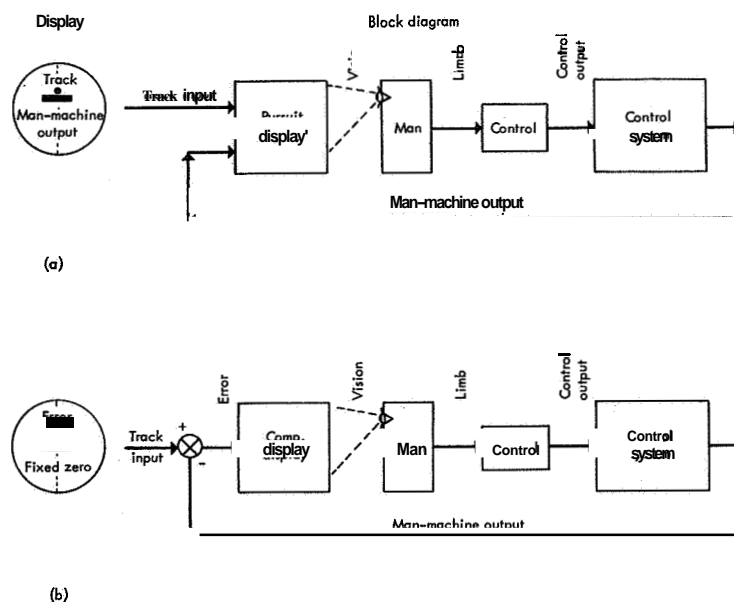


Figure 2.—Pursuit and compensatory tracking. (a) Pursuit; (b) Compensatory.

system output. If he wants to know the present position and rate of the input, he has to compute it by considering the present position and rate of movement of his control in relation to the present position and rate of the error pointer. This computation involves combining kinesthetic and visual information. In order to perform it, one kind of information **or** the other must be recoded.

If the man is able to track with the error held more **or** less at zero, he can learn the statistical properties of the forcing function; but they will be in terms of the control positions and movements which he adopts, not in terms of movements of the display, since they will be minimal. The man's only knowledge of his control system output is also in terms of control positions and movements. Thus if the man can obtain a model of his task, it is likely to be in terms of control positions and movements.

So with a pursuit display the man is likely to have a visual model of his task, while with a compensatory display his model is likely to be kinesthetic. This distinction is important, since it determines the medium in which the modeller can most easily build the model which is intended to mimic the man. I shall begin with the model for pursuit tracking. It is simpler and easier to present since it employs the visual medium.

PERCEIVED POSITION AND RATE INFORMATION

As I pointed out in 1952 (refs. 9 and 10), the man has two sources of information which he can use in attempting to predict the position and rate of the forcing function input half a second ahead. The first source of information is the present position and rate of the input, which a pursuit display shows him directly. Following Gottsdanker (refs. 11 and 12), I do not believe that people can perceive acceleration very well. A brief isolated accelerating **or** decelerating input appears to be apprehended as a constant-velocity input of intermediate rate.

Thus in terms of Bekey's sampled data models (ref. 12) I assume that the rate information used by the man in making a decision is that of Bekey's conventional first-order hold (fig. 3): the average rate since the last decision. From Gottsdanker's results, I believe that the man does not perceive the actual rate at the time of his decision. Thus Bekey's modified first-order hold is probably not appropriate here.

STORED FORCING-FUNCTION INFORMATION

The second source of information which the man can use in prediction is his knowledge of the statistical properties of the forcing-function input. For my present purpose it is enough if I assume that the man learns only the mean frequency and mean amplitude of the input. From the position and direction of movement of the input marker on the pursuit display, the man can predict where the marker will be and the rate it will then have if it moves on the hypothetical forcing function with the mean frequency and amplitude of the actual forcing function.

The man thus has two predictions of the position and rate of the input marker 0.5 second ahead in time. One is based upon position and rate information given directly by the pursuit display. The other is based upon his knowledge of the statistical properties of the forcing-function input. These two predictions are combined in the model by a weighting function.

As far as I can see, all of Bekey's sampled data models except his zero-order hold overshoot whenever the input function reverses its direction of movement. My model also produces overshooting when the input marker reverses direction on an excursion of small amplitude. But my model undershoots a large-amplitude excursion, just as the practiced man tends to undershoot a large-amplitude excursion when the input contains high frequencies. With excursions of intermediate amplitude, the model should fit the forcing-function input

fairly well. The model may then give the impression that it is responding to the present acceleration of the forcing function, just as a practiced man's output may give the impression that he is responding to perceived acceleration.

RESPONSE SELECTION

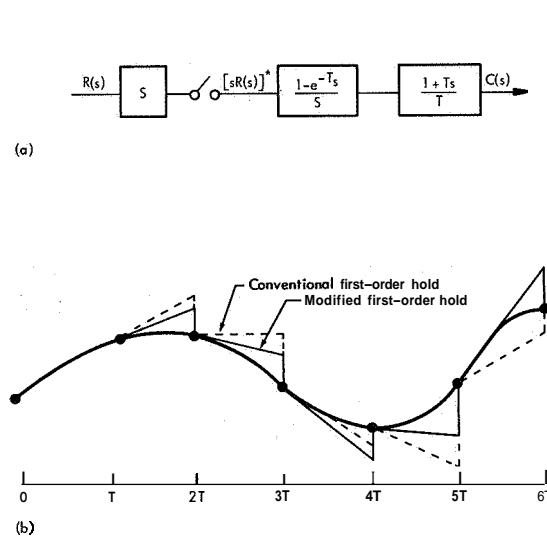


Figure 3.—Modified first-order hold circuit [from ref. 13]. (a) Block diagram; (b) Data reconstruction between samples.

Having computed the position and rate of the input marker 0.5 second ahead, the man faces the problem of how to get his response marker to this position and moving at this rate in 0.5 second. He cannot change his present response during his reaction time, which we can take as about 0.2 second. His problem is thus to move his response marker from where it will be and the rate at which it will be moving in 0.2 second, to where it has to be and to the rate at which it should then be moving 0.3 second later. Clearly, there is a maximum acceleration set by the mass of the part of the limb used in tracking. If the required changes fall inside this limit, there are an infinite number of ways in which the response could be made. The skilled man appears to select the movement which requires the minimum of acceleration and deceleration. This is probably a reasonable assumption to build into the model.

MODEL FOR COMPENSATORY TRACKING

The corresponding model for compensatory tracking uses the movements of the man's control as its basis, not the movements of the display. The role of the hypothetical forcing function with the mean frequency and amplitude of the actual forcing function increases in importance, since the man cannot see the present position and rate of the actual forcing function. I assume that the man makes control movements which he predicts will keep the output of the control system on the hypothetical forcing function and simply modifies these every 0.5 second according to the direction, size, and rate of the error. Because of the disadvantage of the compensatory display, the model involves more computations than are required for the model with the pursuit display.

First, the man has to compute the position and rate of movement of his control a reaction time ahead. Secondly, he has to decide, from its position and direction of movement, where the control will then be in the sequence of to and fro control movements which would generate the hypothetical forcing function with the mean amplitude and frequency of the actual forcing function. Thirdly, he has to compute where the control should move and the rate at which it should be travelling 0.3 second later if it were to continue the to and fro movements which generate the hypothetical forcing function.

So far, I have dealt only with prediction from the stored information about the forcing function. The man also has information from the compensatory display about the direction, size, and rate of his error function. The fourth step is to compute the position and rate of the error marker a reaction time ahead. For this it can be assumed that the marker has a constant rate of movement. The fifth step is to compute where the control should move and

the rate at which it should be travelling in order to bring the error and error rate to zero at the end of the 0.5-second epoch.

As in the pursuit model, a weighting function has to be used to combine the two predictions, one from the stored information about the forcing-function and the other from the displayed position and rate information. The final response selection is similar to that for the pursuit model.

CONCLUSIONS

The main differences between these two models and models based upon servo theory are the following:

- (1) It is aimed to produce data with high-frequency inputs which actually mimics the man, so that human and programmed data are indistinguishable.
- (2) Following Bekey (ref. 13) and Sheridan (refs. 14 and 15) the models are sequential in time rather than continuous.
- (3) The models use statistical information about the forcing function which the man has learned, as well as displayed position and rate information.
- (4) Separate models are given for pursuit and for compensatory tracking, according to the different tasks which the man faces in the two quite distinct situations.

Bekey's sampled data models restrict themselves to using only the position and rate information. Yet they seem to be fairly successful at fitting some aspects of the man's performance with high-frequency inputs, such as the peaks in the error spectral density function in the region of 50 to 100 cycles per minute. I have great hopes that the full models, which use also stored information about the forcing-function input, will fit the man's performance still better.

REFERENCES

1. Stone, M.: Subjective Discrimination as a Statistical Method. *Brit. J. Statist. Psychol.*, vol. 14, 1961, pp. 25-28.
2. Tustin, A.: The Nature of the Operator's Response in Manual Control, and Its Implications for Controller Design. *J. Inst. Elect. Eng.*, vol. 94, 1947, pp. 190-202.
3. Craik, K. J. W.: Theory of the Human Operator in Control Systems. I: The Operator as an Engineering System. *Brit. J. Psychol.*, vol. 38, 1947, pp. 56-61.
4. Craik, K. J. W.: Theory of the Human Operator in Control Systems. II: Man as an Element in a Control System. *Brit. J. Psychol.*, vol. 38, 1948, pp. 142-148.
5. Noble, M.; Fitts, P. M.; and Warren, C. E.: The Frequency Response of Skilled Subjects in a Pursuit Tracking Task. *J. Exp. Psychol.*, vol. 49, 1955, pp. 249-256.
6. Poulton, E. C.: On the Stimulus and Response in Pursuit Tracking. *J. Exp. Psychol.*, vol. 53, 1957, pp. 189-194.
7. Pew, R. W.: The Acquisition of Hierarchical Control Over the Temporal Organization of a Skill. *J. Exp. Psychol.*, vol. 71, 1966, pp. 764-771.
8. Leonard, J. A.; and Newman, R. C.: On the Acquisition and Maintenance of High Speed and High Accuracy in a Keyboard Task. *Ergonomics*, vol. 8, 1965, pp. 281-304.
9. Poulton, E. C.: Perceptual Anticipation in Tracking With Two-Pointer and One-Pointer Displays. *Brit. J. Psychol.*, vol. 43, 1952, pp. 222-229.
10. Poulton, E. C.: The basis of Perceptual Anticipation in Tracking. *Brit. J. Psychol.*, vol. 43, 1952, pp. 295-302.
11. Gottsdanker, R. M.: The Ability of Human Operators To Detect Acceleration of Target Motion. *Psychol. Bull.*, vol. 53, 1956, pp. 477-487.

12. Gottsdanker, R.; Frick, J. W.; and Lockard, R. B.: Identifying The Acceleration of Visual Targets. Brit. J. Psychol., vol. 52, 1961, pp. 31-42.
13. Bekey, C. A.: Sampled Data Models of the Human Operator in a Control System. Rep. 9990-6013-RU-000, Space Technology Laboratories Inc., Los Angeles, Calif., Feb. 1962.
14. Sheridan, T. B.: Some Predictive Characteristics of the Human Controller. Progress in Astronautics and Aeronautics, vol. 13, 1964, pp. 645-663.
15. Sheridan, T. B.: Three Models of Preview Control. IEEE Trans. on Human Factors in Electronics, vol. 7, 1966, pp. 91-102.

19. The Effect of a Random-Sampling Interval on a Sampled-Data Model of the Human Operator'

G. A. Bekey and J. M. Biddle
University of Southern California
A. J. Jacobson
Hughes Aircraft Company

" *
N68 - 15920

There are a number of proposed mathematical models of the human operator in a compensatory tracking function. Some features of a human operator, such as time delay, have been relatively easy to incorporate in most of these models. However, the remnant, or those frequencies in the human operator's output which are not linearly correlated with the input, has never been easy to include in a satisfying, organic fashion in many of the models. In recent years, sampled-data models have been advanced for a number of reasons, one of which is that such models will produce naturally frequencies not present in the input due to the presence of the sampler.

Historically, sampled-data models of the human operator have assumed a periodic sampler. The behavior of such samplers has been studied extensively and certain characteristic features have been well described. For instance, if the model contains such a sampler, the output spectra will have "humps" containing direct and reversed images of the input spectra centered about integral values of the sampling frequency. In addition, there is a null point in the output spectra at each integral value of sampling frequency rather like a deep valley between two hills.

In reference 1 it is stated that carefully analyzed records of actual human-operator outputs failed to show the humped spectra characteristic of a periodic sampler. In that report, the authors theorized that the remnant is due to time-varying parameters, such as gains, time constants, and time delays within the human-operator transfer function. If these parameters may vary, how would the behavior of a human operator model be affected if, assuming it contained a sampler, the sampling interval were allowed to vary in some fashion? Our experiment is an investigation of a sampled-data human-operator model in which the sampling interval varies randomly about some mean value.

EXPERIMENTAL PLAN

We designed the experiment based on a simple analog computer simulation of a sampled-data model of the human operator in a one-dimensional compensatory tracking loop with a constant gain. Two types of data hold, a zero order and a first order, were investigated, and different amounts of randomness were used for a parametric study. Starting with the zero-order hold and a purely periodic sampler, the sampling interval was made to vary in a

*This work was supported in part by the National Aeronautics and Space Administration under Grant NASA-NGR 05-018-022.

Gaussian fashion about the mean interval. The amount of randomness was increased in three steps to a maximum variation of $1\sigma=17$ percent of a mean 0.33-second sampling interval. The experiment was then repeated with a first order hold, and with two successive amounts of randomness with a maximum variation of 12.5 percent of a mean 0.3-second sampling interval. All other model parameters were held fixed for a given case. The model input consisted of the sum of five essentially nonharmonic sine waves. A linear, continuous human-operator model was also constructed and run simultaneously with the sampled-data model. This model was an approximate one, including a gain and lag terms, but without a time delay. This simplification is justifiable in view of the use to which the model was put, as explained subsequently.

Output data were recorded and used as follows. The continuous-model output was recorded and power spectral analyses were performed on the two purely periodic runs. Although useful for comparison, this signal was not of major interest. The sampled model output also was recorded and power spectral analyses were performed mainly for comparison purposes. Although the behavior of this model as the randomness is increased was the primary objective, another signal was used for the reasons given subsequently. This signal was the difference of continuous and sampled model outputs, as shown in figure 1. The time constant and gain (K'' and K') of the continuous model were adjusted so that its frequency response matched that of the sampled model as closely as possible in the low-frequency or input region. This difference signal, when compared with the sampled model output, enhances the higher frequency remnant area and allows scaling in the data reduction to decrease the effect of data channel noise. Due to this fact, our results are based on spectral analysis of this signal. Finally, the sampling-interval statistics were collected to confirm the experimental distribution used and were checked by a chi-squared test to show that they were closely Gaussian.

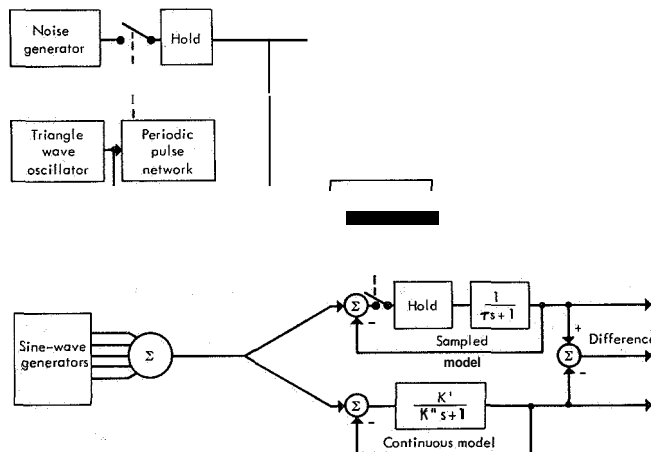


Figure 1.—Block diagram.

EXPERIMENTAL SETUP

The block diagram (fig. 1) shows the general arrangement of the analog simulation. The models and the input sine-wave generators are self-explanatory. The hold on the sampled model, however, must be supplied with a random pulse train, and the remaining blocks in the diagram represent the hardware that generated it. The interval between pulses was required to vary in a Gaussian fashion about a fixed mean. Our method for accomplishing this afforded us the capability of conveniently and independently adjusting both the mean and the variation of the interval between samples.

A triangle wave oscillator was used to provide a stable and symmetrical wave form with a linear slope and equal positive and negative voltage excursions. At each positive and

negative peak of the triangle wave a sample of the signal from the Gaussian noise generator was taken and held. The running slope of the triangle wave was then compared with this sample of the noise in a comparator. The output of this comparator was operated on by a pulse forming network to give a train of clean rectangular samples with a fixed "on" time and a random variation about a mean sampling interval of one half the period of the triangle wave. The statistics of the interval variation were Gaussian and their variance was controlled by the relative amplitude of the rms noise voltage to the peak-to-peak voltage of the triangle wave. The noise voltage was limited to a value no greater than the peak-to-peak voltage to prevent any embarrassing extraneous pulses. The ratio of the limit voltage to rms noise voltage was chosen sufficiently large to have little effect on the interval statistics.

RESULTS AND DISCUSSION

Figure 2 shows the power spectrum of the output of the sampled-data model using a zero-order hold. Note the input frequencies and the "humped" spectra on either side of the 3-cps sampling frequency. This graph, with its strong "valley" trend at 3-cps clearly displays the expected characteristics of a sampled data model. In figure 3 the spectra of the difference signal is shown for the same set of experimental conditions. There is an approximately 10-dB enhancement of the sampling-frequency area over figure 2, the sampled model output. Again, the null at 3 cps is evident. One can easily pick out the peaks which correspond to the direct and reversed images of the input frequencies. Note also the relatively quiet region between the input and sampling frequency regions.

The next three figures (figs. 4 to 6) show the same output, zero-order hold and difference signal for increasing amounts of sampling interval randomness. Histograms of the sampling-interval (AT) statistics are plotted for each case on the corresponding spectrograph. Chi-squared tests on these histograms substantiated the claim that the distributions are Gaussian. These figures show the diminution of the peaks in the sampling region as the randomness increases. The average sampling frequency and the points where spectral peaks would be in the periodic case are marked on each figure. With a AT (1 σ) of only 7 percent we can see a definite decrease. More significant, however, is the "filling in" of the valley at the sampling frequency and also of the normally quiet region between the input frequencies and the sampling area. For the AT (1 σ)=12-percent case (fig. 5) we note a 5- to 10-dB decrease in the peaks, an approximately 7-dB increase in average power level at the valley and a considerable increase in the noise in the formerly quiet region. This trend continues with increasing randomness, and in figure 6 (AT, (1 σ)=17 percent) it is hard to find any resemblance to the spectrum of a sampler, even with the general area marked. Similar results were obtained for the case of model containing a first-order hold (figs. 7 to 9). In these cases the sampling frequency is 3.3 cps, and two degrees of sampling interval randomness, AT (1 σ)=6.5 percent and 12.5 percent, are presented in terms of the difference signal.

CONCLUSIONS

Our experimental results show that:

(1) It might be impossible to determine the presence of an internal sampler by examining the human operator's output spectra if a random sampler is assumed. Noteworthy is the fact that this assumption does not require an unreasonable amount of randomness in the sampling interval.

(2) Assuming a randomly varying sampler our results are compatible with both those presented in reference 1 and those obtained from earlier sampled-data models.

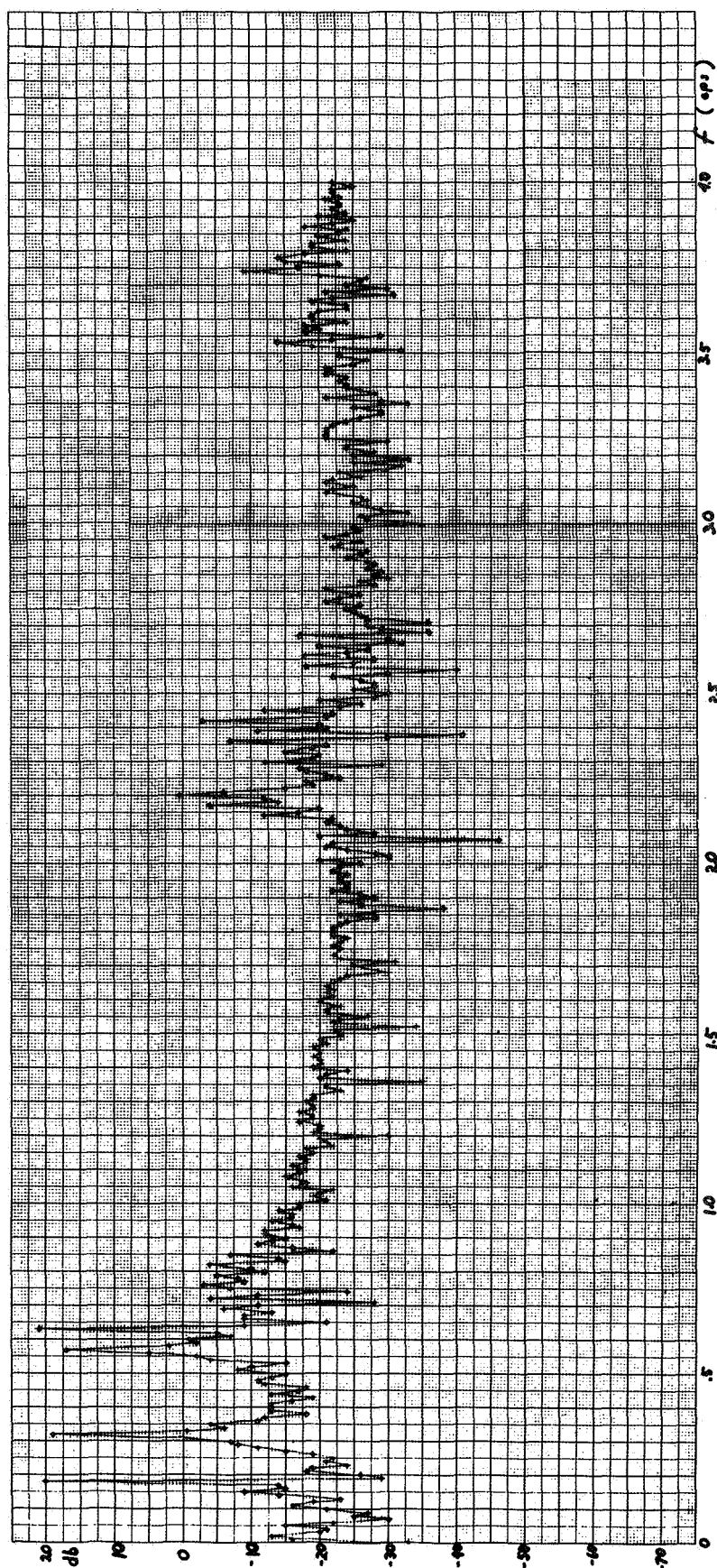


Figure 2 — $Z^{-1}r_0$ -order hold for purely periodic sampled model output.

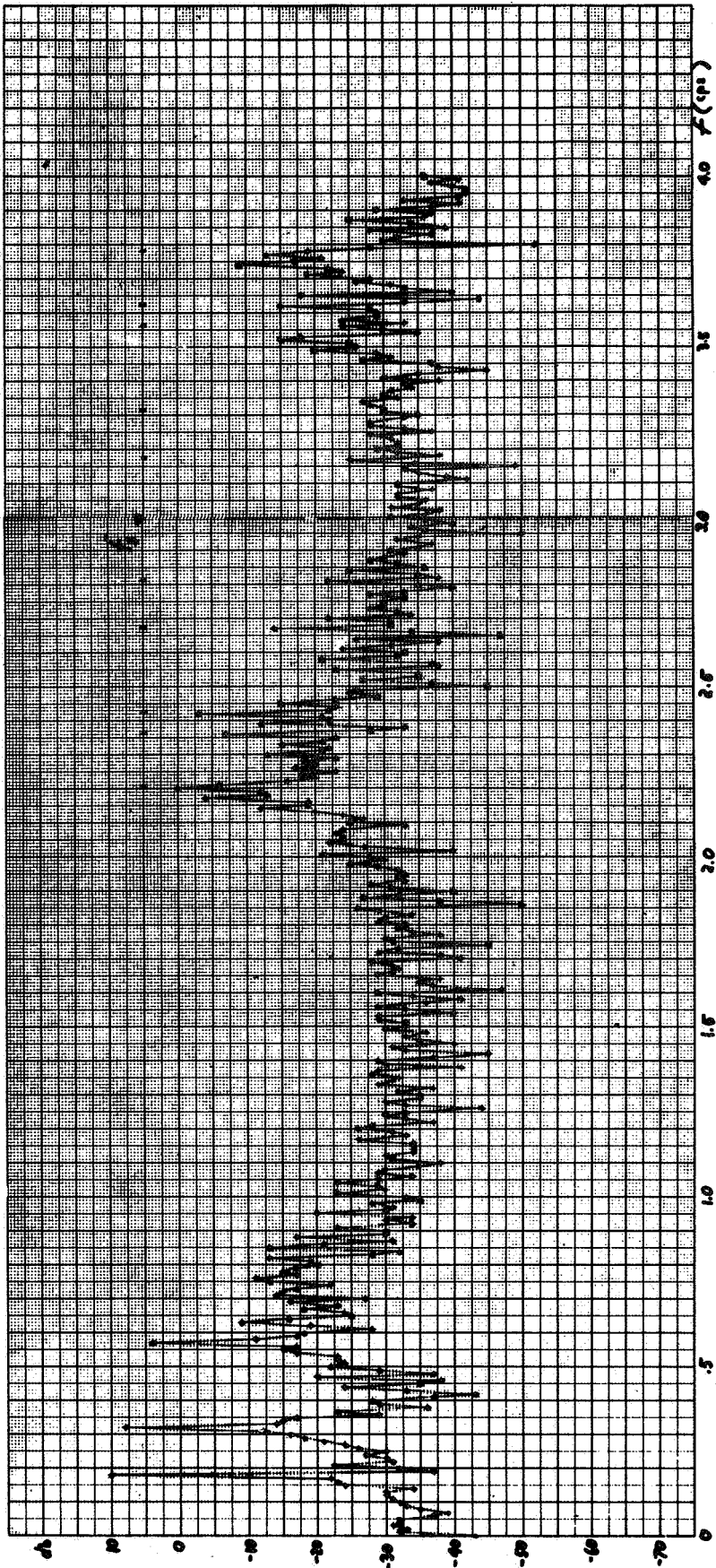


Figure 3 — Zero-order hold for purely periodic difference signal.

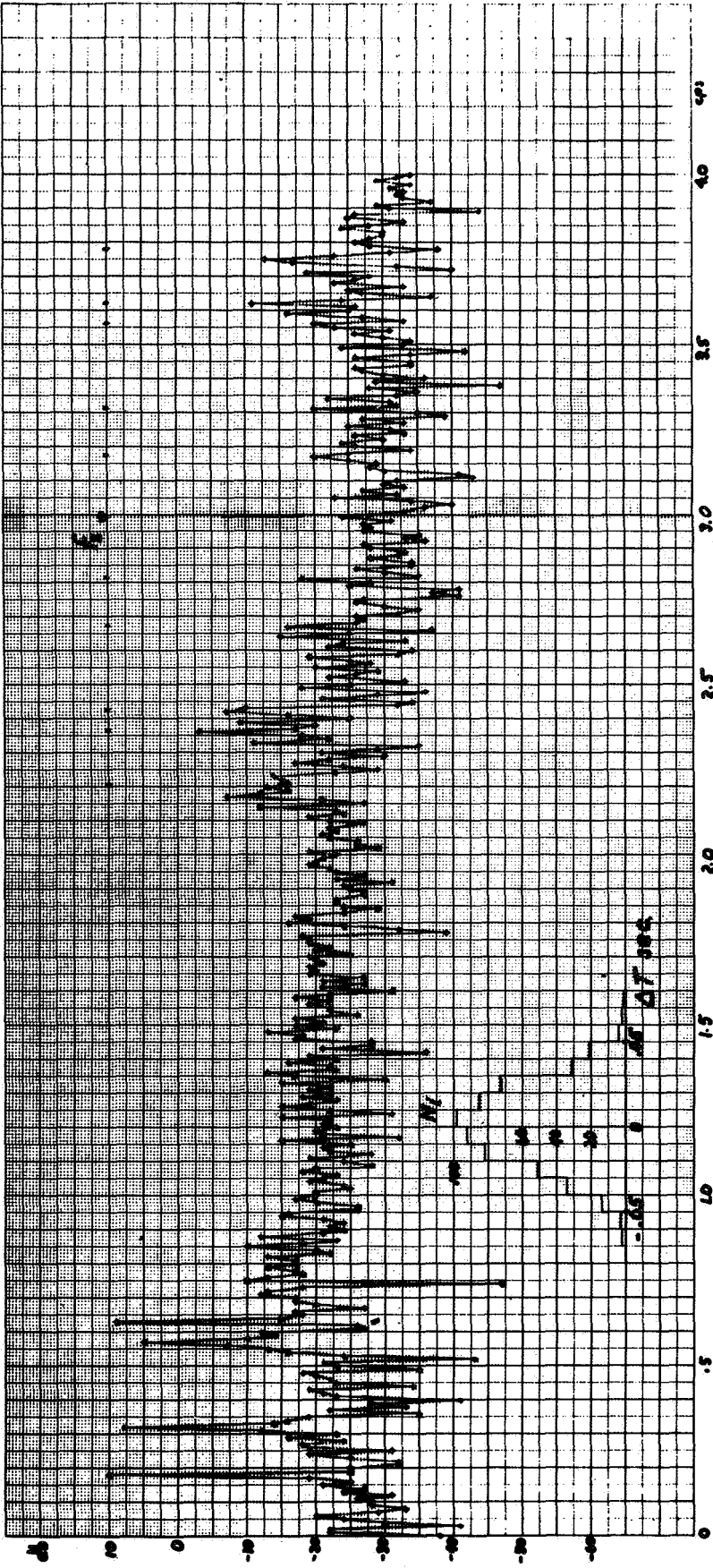


Figure 4.—Zero-order hold for ΔT (1σ)=7 percent and $\tilde{T} \approx 0.33$ second.

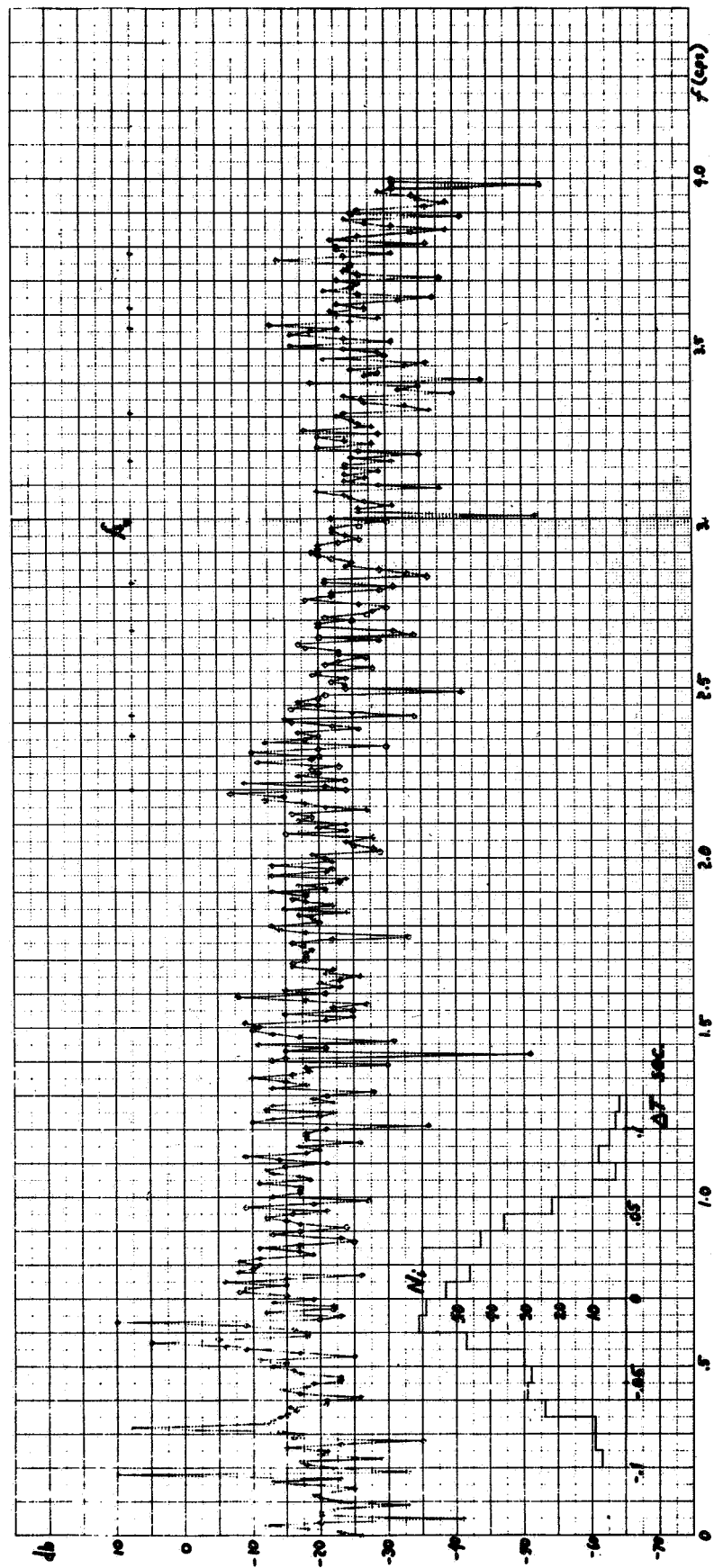


Figure 5.—Zero-order hold for ΔT (1 σ)=12 percent and $\tilde{T} \approx 0.33$ second.

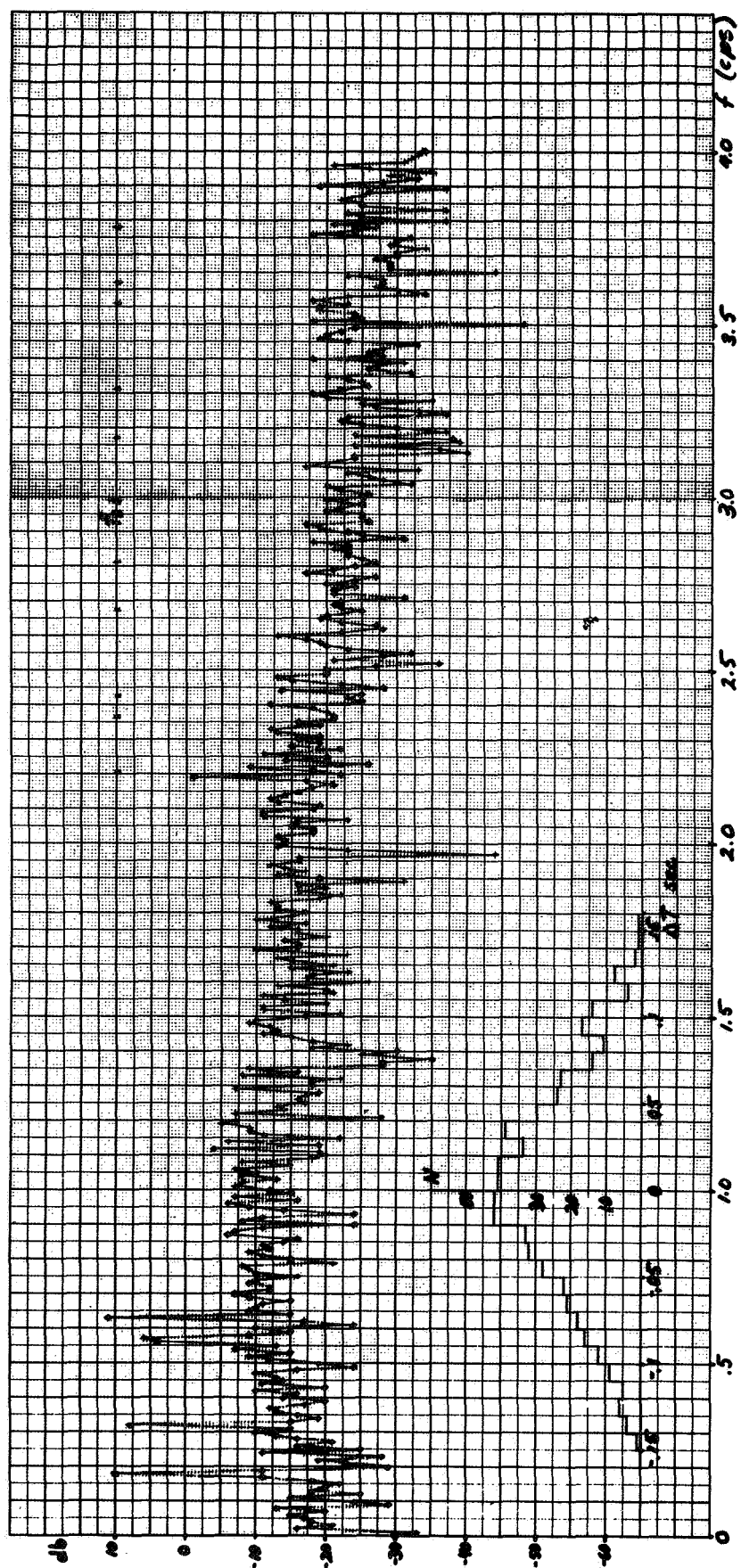


Figure 6.—Zero-order hold for ΔT (1σ) = 17 percent and $\tilde{T} \approx 0.33$ second.

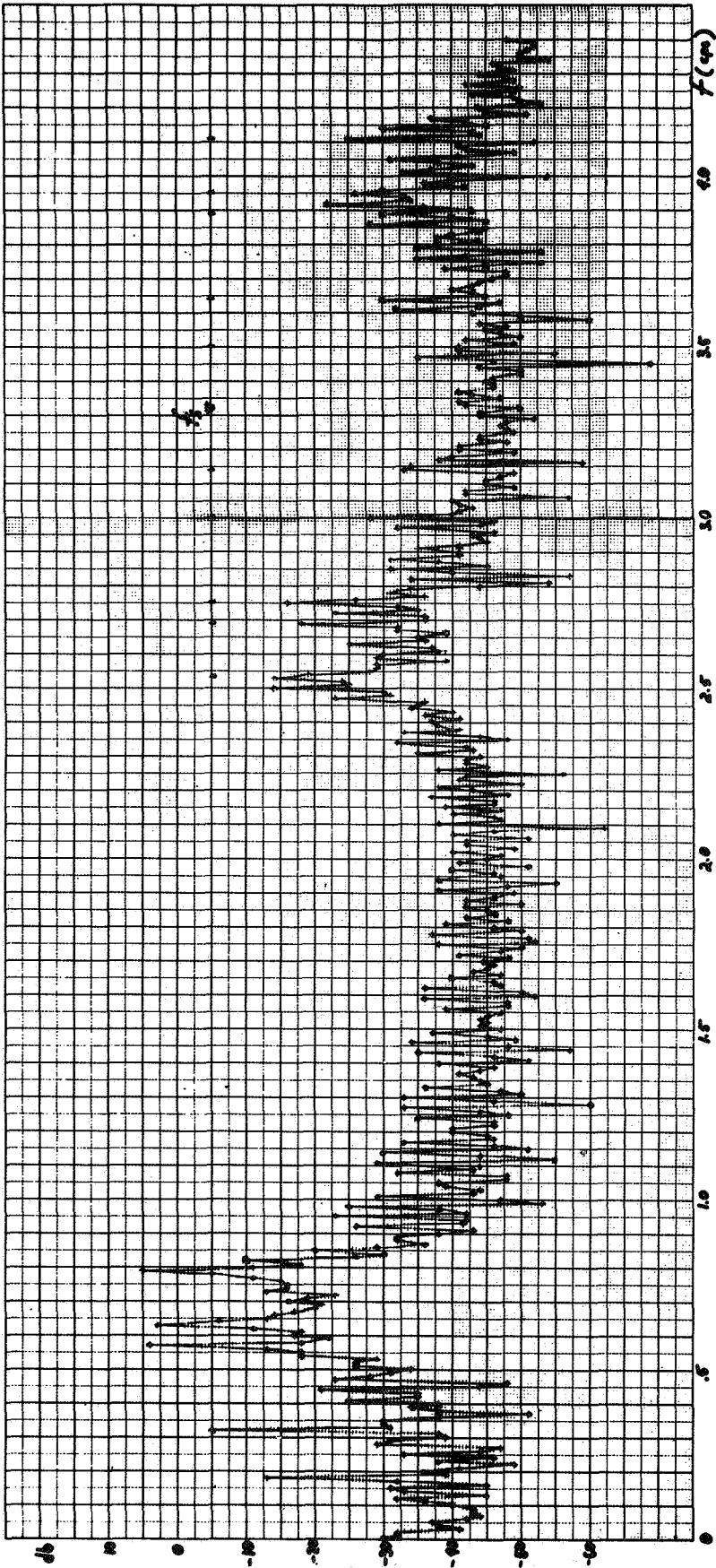


Figure 7.—First-order hold—periodic.

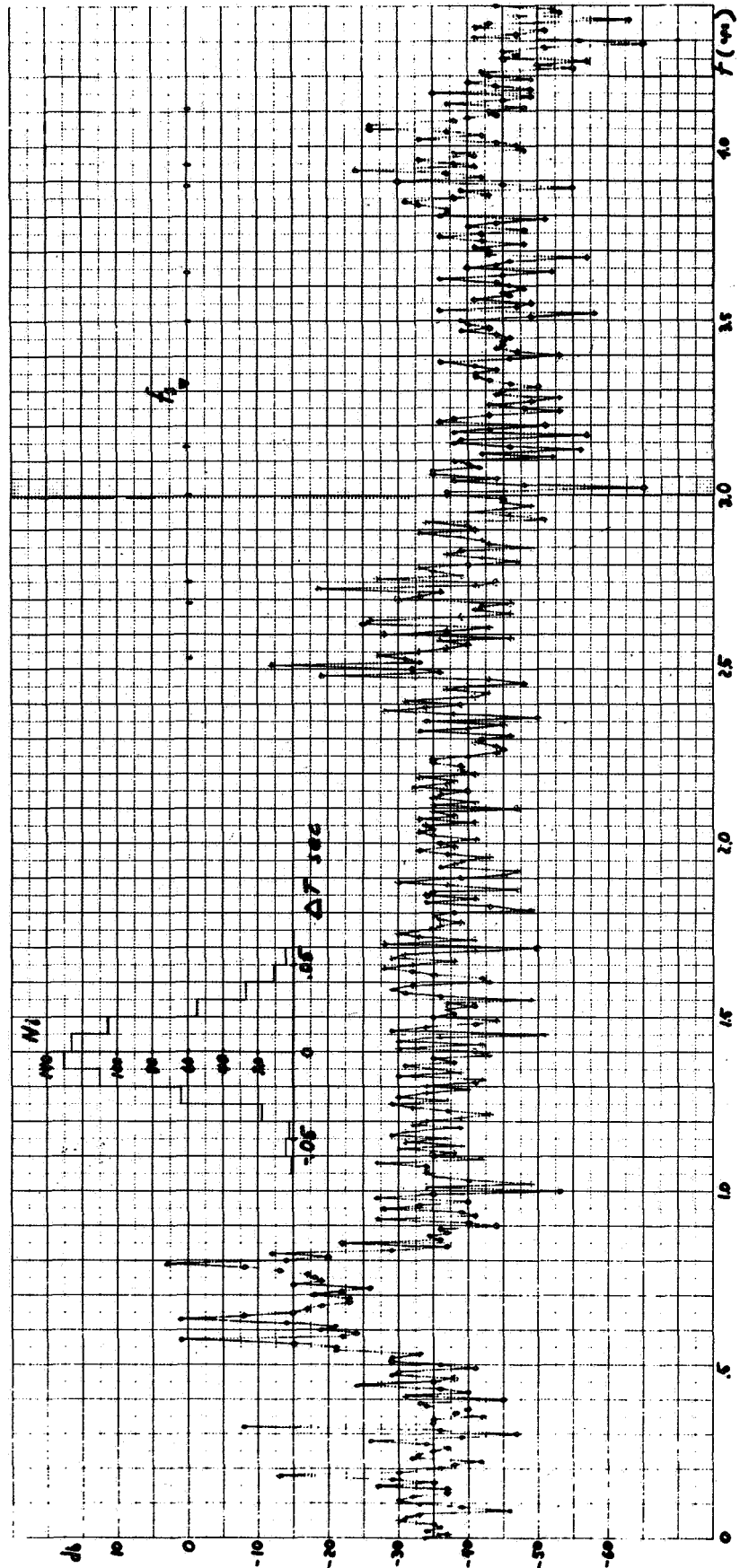


Figure 8.—First-order hold for ΔT (1σ) ≈ 6.5 percent and $\tilde{T} = 0.3$ second.

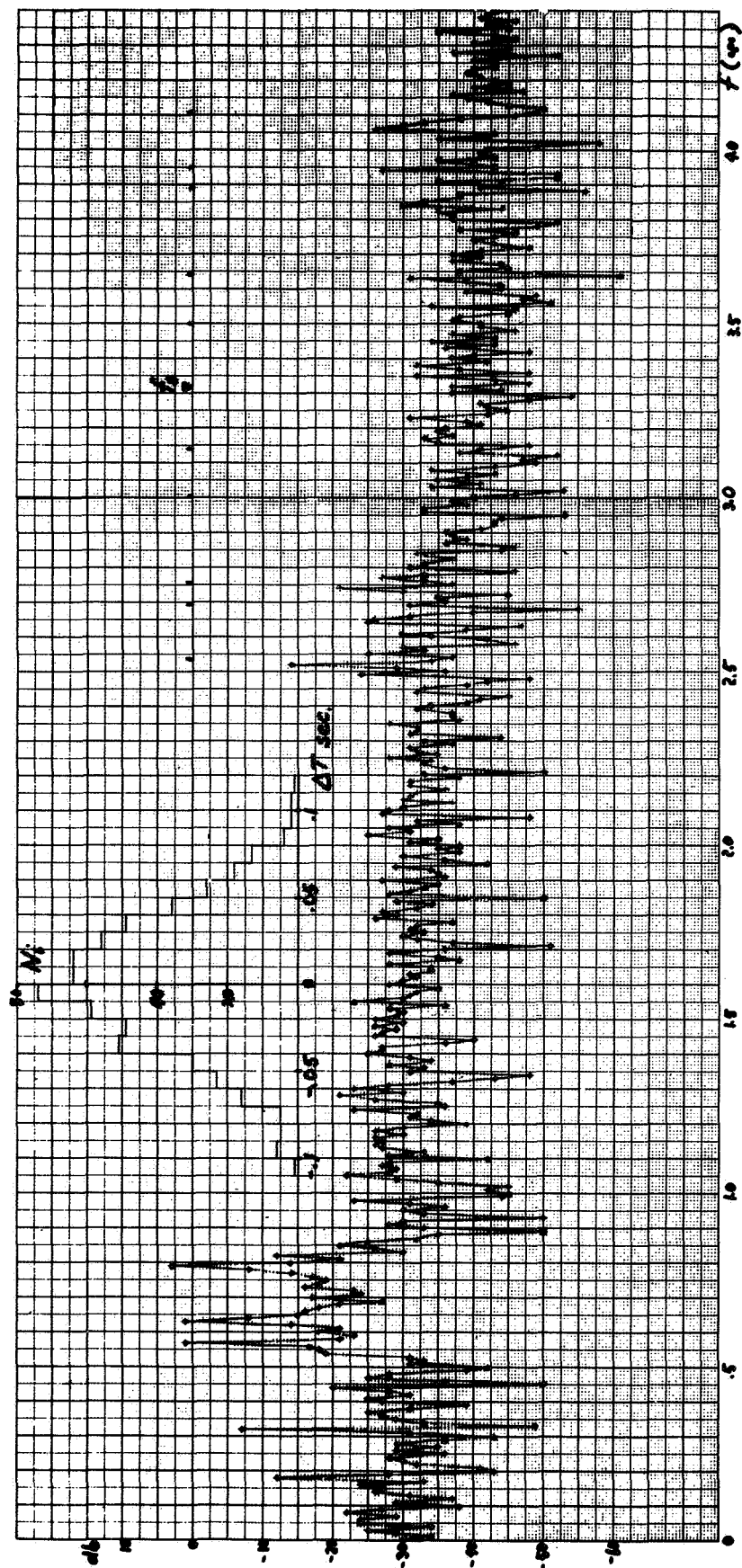


Figure 9.—First-order hold for $\Delta T \cong 12.5$ percent and $\tilde{T} = 0.3$ second.

(3) This then demonstrates the feasibility of using sampled-data models for the human operator, provided only that the sampling interval includes a reasonable amount of randomness.

(4) Finally, our results suggest that attempts at enhancing the remnant portion of the human operator's output and improvement of the signal processing and power spectral analysis will be important in further determination of the validity of sampled-data models for the human operator.

AREAS OF POSSIBLE EXTENSION

We have worked with a random variation in the sampling interval. Furthermore, we used noise with a Gaussian probability distribution to control the randomness. It is not particularly evident that this would be the "best" distribution to use if one were specifically trying to match the model output spectrum to actual human-operator output spectra. In fact, it does not seem unlikely that some sort of deterministic "criterion function" could be used to determine the sampling intervals (in conjunction with some degree of randomness). All these possibilities are in the realm of model matching and form the basis for possible future work.

REFERENCE

1. McRuer, D.; Krendel, E.; Graham, D.; and Reisener, W.: Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15, July 1965.

VII. PHYSIOLOGICAL MODELING

20. Peripheral Versus Central Adaptation: Some Preliminary Results"

N68-15921

*G. C. Agarwal, B. Berman, P. Loehnberg and L. Stark
Presbyterian-St. Luke's Hospital and College of Engineering
University of Illinois*

This paper is devoted to exploring possible decomposition of adaptive mechanisms into those utilizing a variable state of gain α dynamic of the peripheral control system, and those relying upon prefiltering of the preprogrammed, ballistic, intermittent commands which dominate the signal pathways during voluntary rapid movements.

The mechanical-impulse technique is put forward as an experimental means of testing the state of the peripheral control system during various types of adaptive changes. Adaptation that takes place via the peripheral system should be reflected in changes in the mechanical constraints of the hand. The experimental results indicate that in the normal mode the peripheral system remains unchanged for positive and negative mechanical impulses.

The human operator in a control situation is a very versatile, adaptive controller, one that can adapt to many different kinds of changes in operating conditions. Russell (ref. 1), Krendel and McRuer (ref. 2), and Hall (ref. 3) have shown that in the steady state the human operator adjusts his own characteristics. Elkind (ref. 4) has studied the manner in which human-operator control characteristics depend upon the statistical properties of the input forcing function. Sheridan (ref. 5) has examined the response of the process of human adaptation to sudden and gradual changes in controlled element dynamics. Young, Green, Elkind and Kelly (ref. 6) have investigated in some detail the fine structure of the human operator's process of adaptation to sudden changes in system dynamics and input forcing function.

MODEL

A tentative model of the motor coordination system, particularized to the motion of the supinator and pronator muscles of the forearm, is shown in figure 1. The visual loop and the external input are not included. A distinction is made between the postural and voluntary system: a lower level diffuse reflex postural system whose dynamic behavior is controlled via the muscle spindle, and a higher level specific control system which is open-loop with

*

This work was supported in part by the Office of Naval Research under Contract N 00014-67-H-0185.

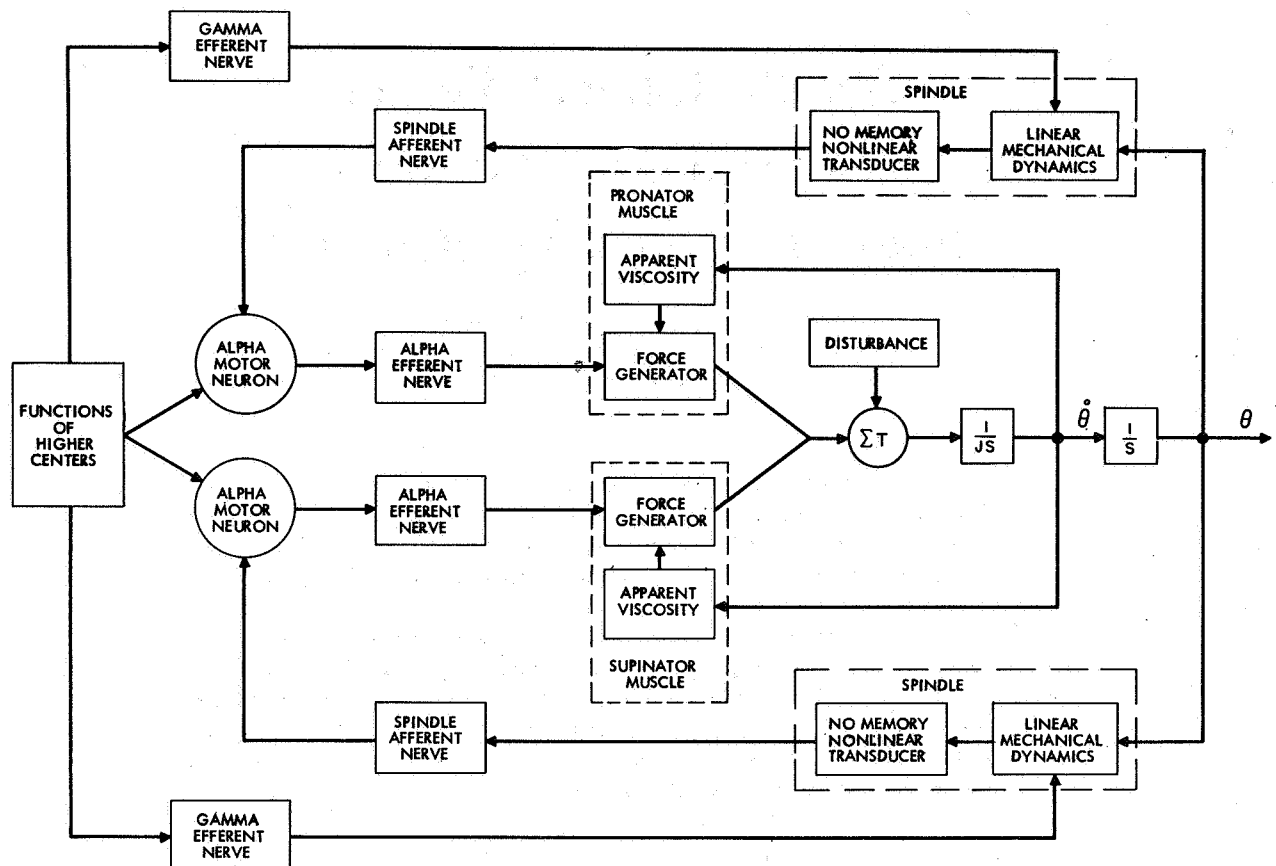


Figure 1.—Motor coordination system.

respect to the proprioceptive feedback during important portions of its behavior. The interaction between these two systems is complex and antagonistic in part.

PERIPHERAL VERSUS CENTRAL ADAPTATION

Various studies by Stark and other investigators (refs. 7 to 9) have shown adaptation in the human operator to take several forms:

(1) Adaptation to a predictable signal which involves recognition of the parameters of the input waveform in a carefully preprogrammed, sequential set of signals to drive the muscles. There is some evidence that proprioceptive sensory feedback is necessary in a physiologically generated form, that is, by active movement, before these "engrams" or adaptive matched filters can be formed. On the other hand, general sensory pattern classification of, for example, visual patterns, may well operate in this manner without requiring active motor participation by the subject.

(2) Adaptation to constraints on random signals has been demonstrated. The discontinuous or sampled-data position control system perhaps here operates in conjunction with higher control of the peripheral servo-loop so that the antagonist muscle brought into action as part of the paired postural servo may act as a damping element.

(3) Adaptation to changes in stick dynamics suggests that alteration of parameters in the postural servo is in fact the mechanism of adaptation. This would certainly be an economical method of adaptive control, requiring only an estimate of "plant" dynamics in the

adaptive alteration of control system compensation. The alternative method of computing control signals in order to provide for compensation of subsequent distortion by the plant seems more complex. In fact, evidence indicating that only a restricted range of compensation can occur suggests indeed that the first scheme, naturally limited by the range of compensatory alterations, is most likely.

We now may make an explicit hypothesis: adaptation to stick dynamics is via the postural loop; adaptation to input waveform is via the direct proprioceptive open-loop path.

EXPERIMENTAL PROCEDURE

To test this hypothesis experimentally a method of measuring gain and dynamic characteristics of the peripheral spindle feedback system is needed. The mechanical impulse is such a test, since it results in a response waveform which defines the system transfer function or the mechanical constants of the hand. Adaptation that takes place via the peripheral system should be reflected in changes in these parameters. Conversely, adaptation entirely in the form of changed input signals should leave the peripheral system unchanged.

Subjects for the experimental test of the hypothesis will be trained or adapted to three sets of conditions: (1) normal mode, (2) altered load or stick dynamics, and (3) predictable waveform. These conditions must be matched for other parameters of the experimental situation. The hypothesis predicts that when the activity of the spindle feedback is now measured by testing as above, only condition (2) will be accompanied by a change in the postural feedback system parameters from their values in the normal mode.

EXPERIMENT

To control the initial conditions of position and velocity, the onset of the torque pulse input, and the mental set of the subject, the subject was instructed to track a specially synthesized triangular waveform displayed on a screen. The triangular waveform synthesized by on-line GE 225 digital computer consisted of a preassigned number of cycles of triangular wave with random slopes after each cycle of a triangular wave of known slopes. The mechanical torque pulse was generated by a torque motor attached to the shaft of the handle which the subject grasped. Positive and negative mechanical pulses were applied at some fixed point on the known triangular wave. The handle position and the torque motor current were recorded and also fed to the digital computer for data processing. The load dynamics was not altered in this experiment.

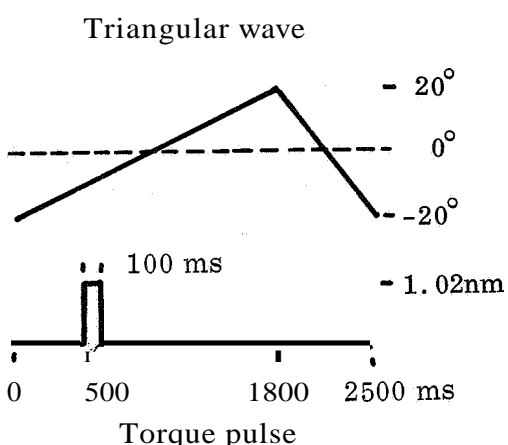

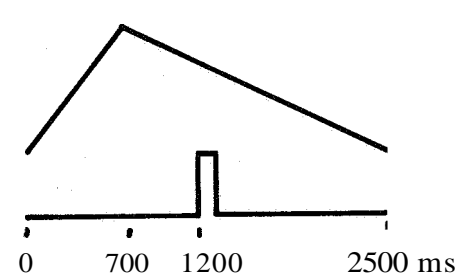
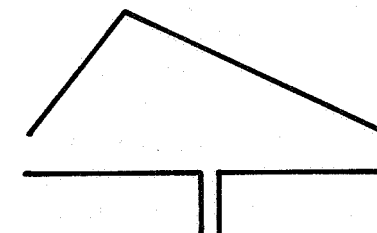
The four possible cases of the direction of rotation of the forearm at the time of the onset of pulse and the direction of the mechanical pulse input are shown in table 1.

RESULTS

It was found that the hand motor coordination system can be described by a second order differential equation with the parameters moment of inertia (J in Kg m^2), coefficient of viscous damping (B in n.m. sec/rad), and torsional spring constant (K in n.m./rad). The values of these parameters for typical responses for the above mentioned cases are given in table 2.

Comparison of the values obtained for moment of inertia under these four conditions indicates that the moment of inertia depends on the direction of forearm rotation at the time of onset of the torque pulse and that the moment is independent of the direction in which the torque pulse is applied. This calculated moment of inertia represents that component of the muscular activity which varies with acceleration, in addition to the moment of inertia of the anatomical structures of the forearm.

TABLE 1.—DIRECTIONS OF FOREARM ROTATION AND MECHANICAL PULSE INPUT

| Case | Form of signal | Rotation | Pulse direction |
|------|--|------------|-----------------|
| | <p>Triangular wave</p>  <p>20°
0°
-20°</p> <p>100 ms
- 1.02nm</p> <p>0 500 1800 2500 ms</p> <p>Torque pulse</p> | Supination | Supination |
| |  | Supination | Pronation |
| |  <p>0 700 1200 2500 ms</p> | Pronation | Supination |
| |  | Pronation | Pronation |

| Case | J, Kg m ² | B, n. m. sec/rad | K, n. m./rad |
|------|----------------------|----------------------|--------------|
| 1 | 1.0×10^{-3} | 1.7×10^{-2} | 2.2 |
| 2 | 1.1×10^{-3} | 0.9×10^{-2} | 2.9 |
| 3 | 0.3×10^{-3} | 0.7×10^{-2} | 2.0 |
| 4 | 0.5×10^{-3} | 0.3×10^{-2} | 2.6 |

DISCUSSION

The experimental results indicate that in the normal mode the peripheral system remains unchanged for positive and negative pulses. A decomposition of the adaptive mechanism suggests the interaction of a variable gain peripheral system responding to stick dynamics and an open-loop system programed by the mental set required for the performance criteria. Further experiments are being planned with altered stick dynamics and task criteria.

REFERENCES

1. Russell, L.: Characteristics of the Human as a Linear Servoelement. M. S. Thesis, M. I. T., 1951.
2. Krendel, E. S.; and McRuer, D. T. : A Servomechanisms Approach to Skill Development. J. Franklin Inst., vol. 24, 1960, p. 269.
3. Hall, I. A. M.: Effect of Controlled Element on the Human Pilot. TR 57-509, WADC, 1957.
4. Elkind, J. I.: Characteristics of Simple Manual Control Systems. Technical Report, Lincoln Laboratory, Mass., 1956.
5. Sheridan, T. B.: Experimental Analysis of Time Variation in the Human Operator Transfer Function. IFAC Congress, Moscow, 1960.
6. Young, L. R.; Green, D. M.; Elkind, J.; and Kelly, J. A.: The Adaptive Dynamics Response Characteristics of the Human Operator in Simple Manual Control. NASA TN D-2255, 1964.
7. Stark, L.; Iida, M.; and Willis, P. A.: Dynamic Characteristics of the Motor Coordination System in Man. Biophys. J., vol. 1, p. 279, 1961.
8. Stark, L.: Neurological Feedback Control Systems. Advances in Bioengineering and Instrumentation, F. Alt, ed., ch. 4, Plenum Press, 1966.
9. Michael, J.: An Analysis of a Predictive Component in the Human Visual Tracking System: M. S. Thesis, McGill University, 1964.

21. State of the Art in Developing Models of Skilled Movement

C. B. Gibbs

National Research Council,
Ottawa, Ontario

N68-15922

Following the Bell (1811) and Magendie (1822) discovery of the proprioceptive pathways, opinion divided sharply on their functions and importance.

Two main theories have emerged in physiological psychology. The "outflow" or "motor" theory of Helmholtz (1867) and his followers asserts that the only factors effective in control are learned patterns of motor volleys to muscles; a released pattern runs its course without proprioceptive influence at high level in the brain. The modern version of the alternative "inflow" or "feedback" theory attributes continuous central functions to proprioception in controlling movement and maintaining perceptual constancy.

Ten young men were tested for their accuracy in making horizontal eye movements toward a visible target and toward the recalled position of a masked target. The latter movements were accurate, and secondary adjustments, made without visual cues, usually reduced the residual errors in primary movements. The subjects were tested both before and after drinking alcohol.

Eight different men moved their right forearm in the vertical plane, without visual cues, to effect alignment with their stationary left forearm. Primary movements and their residual errors were highly variable, but the latter were usually reduced by "blind" secondary adjustments. The subjects were tested for the effects of alcohol consumption and sleep deprivation for 36 hours.

It is argued that the results cannot be reconciled with outflow theory. A model for the control of movement, based on proprioceptive feedback, is outlined.

In this paper, the term "**model**" does not connote the mathematical relations between the input and output of a biological black box but rather, the physiological and psychological mechanisms that are the content of the box. Analogies can be drawn between these mechanisms and mechanical devices which can then serve as models for the biological processes. Ruch (ref. 1) questioned the value of such models when devoid of mathematical treatment, but, in fact, models must usually be developed and refined before appropriate mathematical descriptions can be applied to biological phenomena. In the present state of the art, there has been little advance on Craik's (refs. 2 and 3) model of an intermittent correction servo-mechanism for tracking behavior. In fact, there is no agreement on the best model for even simple, point-to-point movements of limbs or of eyes; controversy on this issue has persisted for over a century.

The "outflow" or "motor" theory of Helmholtz (ref. 4) and his followers in 1867 asserts that the only factors effective in rapid movements are learned patterns of neural impulses in the motor nerves from brain to muscle. A released pattern **runs** its course without proprioceptive influence at high level in the brain.

The alternative "inflow" or "feedback" theory of control has persisted in one form or another since the discovery of the proprioceptive, feedback pathways by Bell (ref. 5) in 1811 and Magendie (ref. 6) in 1822. In the modern version of the theory, important functions are attributed to proprioception at high level in the brain, in maintaining perceptual constancy (ref. 7), and in the continuous monitoring of rapid movements (ref. 8). There are difficulties for this view, as pointed out by Ruch (ref. 1): "Its protagonists have been unable to envisage the precise manner in which proprioception is used in the control of movement."

Craik (refs. 2 and 3) and Vince (ref. 9) confirmed Woodworth's (ref. 10) observation that rapid movements of limbs are not under continuous control by vision, and they found that proprioceptive reaction time was little shorter than visual reaction time. Together with most other workers in the field, Craik supposed that this long latency precluded continuous proprioceptive control of rapid movements.

Other objections that are often made to inflow theory are based on the alleged, gross inaccuracy of proprioceptive information and low awareness of these sensory data. Ludvigh (ref. 11) showed that subjects tested in darkness may not be aware that they have made eye movements as large as 10° .

The requirements for establishing inflow theory are:

(1) To envisage a control method or model that permits continuous proprioceptive involvement at high level in the brain in the control of rapid movements, despite known transmission latencies in the motor-proprioceptor circle of nerves.

(2) To use the model to predict the outcome of experiments that are crucial to the two control theories.

(3) The experiments should show that the accuracy of proprioception is much higher than is suggested by the data of Ludvigh (ref. 11) and others.

In respect to the first requirement, it has been pointed out (ref. 12), that there are lawful relations between the direction of movement and the discharge of specific groups of proprioceptors (ref. 13). The speed of movement is closely related to the rate of change of frequency in the discharge of primary, proprioceptive nerve endings (ref. 14), and this discharge is conducted very rapidly to the cerebellum. The speed and direction of movement could therefore be controlled directly by proprioceptive feedback, once the relevant relations were learned. The extent of movement could be determined by integrating the speed signals from proprioceptors over time to provide integral-error control (ref. 8).

Integral-error control would be based upon judgments of the speed of movement (kinaesthesia) and of the duration of a primary adjustment of about 0.25 second, which is the period of integration. Two sources of error therefore exist, but subsequent adjustments based on the resting, proprioceptive discharge (statokinesis) have only one source of error and should therefore reduce any residual overshoot or undershoot in the primary adjustment.

Outflow theory makes no provision for sensing the amplitude and direction of residual error in a primary movement, and in this view, there could be no adaptive, secondary corrections to errors in these movements.

EXPERIMENT 1: THE ACCURACY OF EYE MOVEMENTS OF ACQUISITION AND RECALL, WITH AND WITHOUT ALCOHOLIC STRESS

METHOD.—Electro-oculographic records were made, using an Offner Dynograph recorder, of eye movements of acquisition and recall, spanning 1° , 2° , 4° , or 8° in the horizontal plane. White cards, with black target crosses, were placed 28.7 inches from subjects' eyes and 5 feet from a 25-watt lamp bulb. The seated subject used a dental bite to steady his head and first fixated the left cross until the experimenter ordered an eye

movement to the right cross. The left cross was covered by a white, irregular mask and subject was then told to direct his gaze toward the recalled, left position. Following the primary movement of recall and subsequent "blind" adjustments, the mask was removed from the left target and the subject was told to fixate this cross precisely to remove any error in the previous adjustments. The subjects were immediately given detailed information on their final error before the mask was removed.

Ten men between the ages of 21 and 36 years, were given 10 of these cyclic runs on each of the four angles presented in irregular order, on each of 3 days. On the fourth day, the subjects had 10 runs on each angle before consuming alcohol. The subjects then drank whiskey, rum, or gin, diluted to their taste. Breathalyzer readings were taken as required, and the subjects had 10 more runs on each of four angles when their breathalyzer readings were about 0.05 percent and again when further drinking raised their breathalyzer readings to nearly 0.1 percent.

The above procedure had the advantage that an ophthalmograph could be used to check the relatively new electro-oculographic method of recording, and 400 runs were recorded simultaneously by the two methods. Care was taken to eliminate visual reference points, that is "background anchors" of the left, recalled position. Target positions were varied every 10 runs within a 2° zone, centered about a point 3° left of the zero position directly ahead of subject. As a further check, 400 control movements of both acquisition and recall were made in a darkened room. Two small lights were switched on or off, to provide a target or recall position, in cyclic runs of the type described above.

RESULTS.—The data were similar from the ophthalmograph, the electro-oculographic control runs using small lights in a darkened room, and the electro-oculographic runs in light. The latter data, from fourth day are presented and discussed.

There were large individual differences in "proprioceptive sensitivity," which is defined as the smallest residual error in primary movements of recall that evoked a "blind," directionally correct amendment on 50 percent of the occasions when the error appeared. This limit lay between 0.25° and 0.50° for four subjects, between 0.5° and 1° for five subjects, and above 1° for one subject. The attempted 1° movements of the last subject were often obscured on the electro-oculographic chart by drifts and tremor in the records. The eye movements of recall made by seven subjects were accurate and consistent from the beginning of practice, but three subjects made variable and inaccurate recall movements at first. They showed considerable improvement from practice with immediate knowledge of results.

Drift and tremor in the records set limits to the measurement of a subject's "proprioceptive sensitivity" after he was trained. In 7 of the 10 subjects, oscillations of the recording pen increased greatly after drinking had raised their breathalyzer level to about 0.1 percent. Eye movements between the crosses separated by 1° could not be distinguished from the background noise in the records of these subjects, and the results at this angle are not presented in detail.

The best available estimate of an accurate eye movement was the mean of each subject's 10 movements toward a visible target after fine, secondary movements of fixation were completed. This mean was taken as a standard and compared in turn with the mean amplitude in the other three phases of movement. The differences from the standard were the mean errors of subjects, and these were the data treated in the analysis of variance shown in table 1. The errors of all subjects were totalled irrespective of sign (undershoot or overshoot) and used to compute the group, arithmetic means given below.

In rapid, primary movements toward a visible target, the mean group error (difference from standard) was 0.54° , and it was 0.69° in primary movements toward the recalled position of the masked target. The difference was not statistically significant and could have

TABLE 1.—ANALYSIS OF VARIANCE OF ERRORS IN DIFFERENT PHASES
OF MOVEMENT, ANGLES AND BREATHALYZER LEVELS
[The standard deviation of the errors, computed from the residual is 1.17°]

| Source | Df | Mean square | F ratio |
|-------------------------|----|-------------|--------------------|
| Subjects (S) | 9 | 6.687 | ^a 4.86 |
| Breathalyzer levels (B) | 2 | 29.707 | ^a 21.59 |
| S × B | 18 | 2.388 | 1.74 |
| Angles (A) | 2 | 67.951 | ^a 49.38 |
| A × B | 4 | 3.775 | ^b 2.73 |
| A × S | 18 | 2.285 | 1.66 |
| A × S × B | 36 | 1.665 | 1.21 |
| Phases (P) | 2 | 19.514 | ^a 14.18 |
| B × P | 4 | 0.672 | |
| P × S | 18 | 5.196 | ^a 3.78 |
| P × S × B | 36 | 2.713 | ^a 1.97 |
| P × A | 4 | 7.002 | ^a 5.13 |
| P × A × B | 8 | 2.159 | 1.57 |
| P × S × A | 36 | 1.636 | 1.19 |
| S × P × A × B | 72 | 1.376 | |

^aSignificant at the $p < 0.01$ level.

^bSignificant at the $p < 0.05$ level.

arisen by chance. Secondary, "blind" corrections, with the left target masked, reduced the primary, recall error from 0.69' to 0.40°, and this reduction was statistically significant ($p < 0.01$). At breathalyzer levels of 0, 0.05 percent and 0.1 percent, group mean errors were 0.38°, 0.52°, and 0.74°, respectively; with one exception, subjects produced their largest errors at the highest breathalyzer level.

The mean, group errors were 0.28°, 0.53°, and 0.83° at angles of 2°, 4°, and 8°, respectively, or 14, 12, and 10 percent expressed as a percentage of the actual target angle.

Table 1 shows that some interaction effects were significant. An increase of breathalyzer level did not have equal effects at all angles; errors in the three phases of movement were in different ranking with different subjects, and at different angles. Interaction effects were therefore tested by X^2 tests, based upon the number of subjects who produced conflicting results in different conditions. For example, the "blind" secondary corrections to recall movements reduced mean residual errors in nine subjects but the other subject increased errors ($X^2=6.4$; $p \approx 0.01$). The interactions do not invalidate the significant differences ($p < 0.01$) between main factors.

Following primary recall movements, subsequent "blind" adjustments were classified as correct **or** incorrect, depending on whether they had decreased **or** increased the residual error at a time 1.25 seconds after the primary recall movement began. Table 2 shows the numbers of correct and incorrect "blind" adjustments at three angles and breathalyzer levels, and also their ratios and X^2 values.

| | 2° | | 4° | | 8° | |
|----------------------------|-------|----|------|----|------|----|
| | C | I | C | I | C | I |
| Breathalyzer=0 percent: | | | | | | |
| Number | 51 | 16 | 59 | 21 | 75 | 10 |
| Ratio | 3.2: | 1 | 2.8: | 1 | 7.5: | 1 |
| X^2 | 21 | | 24 | | 50 | |
| Breathalyzer=0.05 percent: | | | | | | |
| Number | 43 | 15 | 60 | 19 | 73 | 16 |
| Ratio | 2.9: | 1 | 3.2: | 1 | 4.6: | 1 |
| X^2 | 13.4 | | 21 | | 36 | |
| Breathalyzer=0.10 percent: | | | | | | |
| Number | 24 | 13 | 58 | 23 | 52 | 27 |
| Ratio | 1.85: | 1 | 2.5: | 1 | 1.9: | 1 |
| X^2 | 3.2 | | 15 | | 8 | |

On nondrinking runs, the ratio of correct to incorrect adjustments was highest with the 8° target, because the absolute size of residual error varied directly with the target angle. As dosage increased, there was an increase in the size of absolute errors, but a sharp decline in the ratio of correct to incorrect adjustments. The ratio was not significant at the 2° angle, with a breathalyzer reading of 0.1 percent, but all other X^2 values were significant ($p < 0.01$).

EXPERIMENT 2: THE ACCURACY OF ALINING THE TWO ARMS WITHOUT VISUAL CUES

METHOD.—A seat was provided with a lever on both the left and right sides. The seat and lever pivots were made adjustable so that the pivots could be closely aligned with the elbow joints of the subjects, who were seated with their upper arms vertical and forearms pointing directly ahead. The levers could also be adjusted to the length of a subject's forearm when he grasped a handgrip provided at the far end of the lever. The levers lay beneath the subject's forearms and short channels on the levers supported the subject's elbows. Vertical movements of the forearm, pivoting about the elbow joint, imparted almost identical, rotary movements to the lever because of the close similarity in lengths and the positions of the pivots of lever and forearm. The left lever could be fixed in any one of three positions by the use of a spring-loaded plunger. In the "middle" position, the subject's left upper arm was vertical, with the forearm horizontal, pointing ahead. In the "up" position, the left forearm and lever were pivoted upward to form an angle 30° above the horizontal, middle

position. The "down" position was formed by rotating the forearm and lever downward to form an angle 30° below the horizontal, middle position. Movements of the right lever rotated a potentiometer which imparted movements to the pen of a moving-chart recorder. The equipment was calibrated to permit analysis of the extent of rotary movements of the right forearm, differences or errors between the positions of the right and left forearms and the time taken for movements.

Eight young men were tested for their ability to move their right forearm into alignment with their stationary left forearm. The right forearm made rotary movements, pivoting about the elbow joint in the vertical plane. In each test session, the subject had one trial in each of the three positions, presented in different order in the various sessions. The subjects started each run in the up and the down positions with the right forearm below and above the left forearm respectively, to form an initial angle of about 70° between the forearms. In the middle position the right forearm started above and below the left forearm with similar frequency. At an order from the experimenter, the subject made a primary movement of the right forearm, with his eyes closed, in an attempt to align the forearms. The subject was allowed to make any desired number of subsequent adjustments until he was satisfied that the two forearms were aligned. He was given immediate knowledge of his final error.

As in the previous study of eye movements a major question was whether residual errors in rapid, primary movements would be reduced by subsequent adjustments, but the second experiment had a further purpose in studying the interaction between moderate sleep deprivation and consumption of alcohol. The subjects were kept awake for 36 hours, starting at 8:00 a.m. on the first day and ending at 8:00 p.m. on the second day. During this period, each subject had eight nondrinking practice sessions, comprising one trial at each of the three positions for the stationary left forearm. On the first day, each subject had run 1 between 8:00 and 10:00 a.m., run 2 between 1:00 and 3:00 p.m., run 3 between 5:00 and 8:00 p.m., run 4 between 10:00 and 12:00 p.m. On the second day, run 5 was taken between 1:00 and 3:00 a.m., run 6 between 6:00 and 8:00 a.m., run 7 between 10:00 and 12:00 a.m., run 8 between 1:00 and 3:00 p.m.

The subjects had their ninth and final run between 4:00 and 6:00 p.m. after drinking whiskey or rum. The average breathalyzer reading was 0.6 percent for the group during the final run.

RESULTS.—Comparisons were made between the size of the residual error at the end of primary movements and the error that remained when subject had completed subsequent adjustments. The adjustment was marked as correct or incorrect when it reduced or increased the initial error in the primary movement. The numbers of correct and incorrect movements are shown in table 3 for the three positions of the left forearm and for the different breathalyzer levels.

TABLE 3.—NUMBER OF CORRECT (C) AND INCORRECT (I) SECONDARY ADJUSTMENTS AND ARITHMETIC MEAN ERRORS (E) AT END OF RUN

| Breathalyzer,
percent | Position | | | | | | | | | Total | | |
|--------------------------|----------|---|--------|--------|---|--------|------|----|--------|-------|----|--------|
| | Up | | | Middle | | | Down | | | | | |
| | C | I | E, deg | C | I | E, deg | C | I | E, deg | C | I | E, deg |
| 0 | 52 | 5 | 3.8 | 47 | 8 | 5.1 | 41 | 12 | 3.6 | 140 | 25 | 4.2 |
| 0.06 | 5 | 0 | 2.4 | 7 | 0 | 6.4 | 6 | 1 | 3.4 | 18 | 1 | 4.1 |

The errors of undershoot and overshoot that remained after subject stopped making adjustments were totalled regardless of sign (positive for overshoot and negative for undershoot) and arithmetic means were computed for each position and breathalyzer level. These arithmetic means are also shown in table 3. The standard deviations of errors were 3.51° and 3.34° on nondrinking and drinking trials, respectively.

A total of 140 adjustments reduced the residual error in primary movements and only 25 increased such error ($\chi^2 \approx 77$, $p < 0.01$).

Run 1 on the first day was taken between the hours of 8:00 and 10:00 a.m., and the mean arithmetic error for the group was $4.5''$. On run 5, taken between 1:00 and 3:00 a.m. on the second day, the mean group error was 4° , a slight reduction of error, contrary to the usual finding of reduced efficiency at night. Five subjects had smaller mean error on run 5 than on run 1, and three subjects showed the opposite result. Run 8 was taken between 1:00 and 3:00 p.m. on the second day when the subjects had been without sleep for about 30 hours, but the group mean error remained the same as on run 1, at 4.5° . Comparison of each subject's errors on run 1 and run 8 showed that four subjects were better and four subjects were worse on the former than the latter run. It appeared that 30 hours of sleep deprivation had not significantly reduced the accuracy in the group as a whole.

There were only 3 hours difference in sleep deprivation on runs 8 and 9; the main difference between runs was that the subjects had consumed alcohol before run 9. Nevertheless, the mean group error was 4° for run 9, compared with $4.5''$ for run 8. On an individual basis, the errors of five subjects were lower on the drinking run 9, than on the nondrinking run 8; three subjects showed the opposite result.

In run 1, the subjects were sober and fully rested but the group mean error was 4.5° , as compared with a mean error of 4° on run 9, when the subjects had been deprived of sleep for 33 hours and alcohol consumption had produced an average breathalyzer reading of 0.06 percent. Individual comparisons showed that the errors of four subjects were lower on run 9 than on run 1, two subjects had equal scores on both runs, and only two subjects performed better on run 1 than on run 9.

On each run, each of the eight subjects had one trial in each of three different positions, so that 24 matched comparisons could be made between the errors made on two different runs. Chi-square tests showed no significant difference in accuracy when comparisons were made between the following runs: run 1 and run 8, run 1 and run 9, and run 8 and run 9. There was no significant change in accuracy as a result of sleep deprivation and/or alcohol consumption either singly or alone, in the experimental conditions.

Comparisons were made between the times taken in each trial before each subject was satisfied that his secondary adjustments were accurate. In comparing run 1 with run 8 and run 1 with run 9, 18 trials took longer on the later runs, 8 and 9, than on run 1 and the opposite result appeared in only 6 trials. The value of χ^2 was 6, near the 6.64 value needed for significance at the $p=0.01$ level. In comparing the last nondrinking run 8 with the drinking run 9, there was no significant difference in times. The group mean times and standard deviations were, respectively, 8.34 seconds and 3.91 seconds, on nondrinking runs, and 9.48 seconds and 4.62 seconds on the drinking run. The range of times in the nondrinking runs was from 1.7 to 20.9 seconds and in the drinking run was 3.5 to 23.1 seconds.

It may be concluded with a high degree of probability that sleep deprivation for 30 hours produced a significant increase in the time taken to reach a decision, but not in the precision of executing a movement. Alcohol consumption that produced a breathalyzer reading of about 0.06 percent, plus a further 3 hours without sleep, caused no significant change in performance between runs 8 and 9.

Arithmetic means give a better estimate than algebraic means of absolute errors because an equal number of large errors of overshoot (+) and undershoot (-) can cancel each

other in computing the algebraic mean which will then suggest a high, but spurious, level of accuracy and consistency. However, the arithmetic mean does not indicate whether or not movement and errors are centered on the target, and algebraic means were also computed to determine the central tendency. When errors are either predominantly positive or negative, there is little difference between arithmetic and algebraic means, but a mixture of overshooting and undershooting will reduce the algebraic mean as compared with the arithmetic mean.

There was considerable variation in individual performance; the ranges of subject's mean errors were from 3° to 6° arithmetic mean and from $+0.5^{\circ}$ to -6° in the algebraic mean. The errors of six subjects were predominantly due to undershoot; two subjects had the opposite tendency and these had the lowest arithmetic mean errors of 3.0° and 3.5° as compared with 4.4° for the other six subjects.

In the group as a whole, on nondrinking runs, the arithmetic and algebraic means were 4.2° and 1.8° , respectively. The extent and residual errors of primary movement were highly variable; about 20 percent of these movements had errors of 20° or more, but in 32 percent of cases the residual errors were 4° or less. There was similar variation in the number of adjustments that were made following a primary movement; no adjustment or only one adjustment was made in 41 percent of cases, whereas seven or more adjustments were made to 15 percent of primary movements.

DISCUSSION

A previous study of eye-movements (ref. 14) produced a mean eye error and standard deviation of 0.28° and 1.4° , respectively. These figures are close to the 0.38° eye error and 1.17° standard deviation of the present experiments although the studies differed widely in conditions and techniques of measurement. In the present study, the small errors that remained after "blind," recall adjustments were completed might not be functionally important in seeing the target clearly. When fixation is attempted, there are frequent flicks to correct drift of the eyes, so that a target image travels around a central area of the fovea approximately 0.36° in diameter (ref. 15). The zone of clear, foveal vision is approximately 3° , and relatively gross, saccadic movements of the eyes achieve adequate acuity for reading. In the nondrinking runs of this study, few corrections were made to recall movements when the target mask was removed and the target was seen. Well-practiced subjects made adjustments of 0.5° or more on only 3.5 percent of the trials. In 15 percent of the trials, smaller corrections could be detected in the records.

The above data are compatible with the hypotheses that proprioception is used to control primary eye movements and fine, secondary adjustments toward either a visible target or the recalled position of a target. The alleged inaccuracy of proprioception has often been cited as a crucial objection to any important central role of this modality in movement control. The precision now shown provides a factual basis for dismissing this venerable objection. The underlying fallacies in experiments and interpretations must also be disposed of because they have been repeated, in one form or another, for more than a century. In a typical case, Ludvigh (ref. 11) states that proprioceptors in extraocular muscles cannot position the eye with higher precision than $10'$. His subjects fixated a small light in a darkened room. This light was extinguished but soon replaced by another, which was displaced from the first by varying amounts. The subjects fixated the second light and reported whether it was displaced to the left or the right of the first. Their judgments were not reliable for displacements less than 10° , in contrast to the present study where most subjects could reliably detect and correct an error of 1° or less. In the Ludvigh study, the subjects made verbal reports of the minimum sensation which reached the level of consciousness. The present study determined

the minimum data that are effective in control, regardless of the irrelevant issue of the subject's awareness of the information. Relatively few of the proprioceptive pathways project to the cerebral cortex, the neural center of consciousness. It has been pointed out previously (ref. 12) that a major advantage of delegating control to proprioception is to release the limited span of conscious attention **for** its primary function of detecting environmental changes of biological importance. In common experience, a long established skill is often disrupted when attention is paid to detailed execution. Despite the weight of scientific and experimental observations, many have discounted feedback theory, mainly because low awareness of proprioceptive data often leads to inaccurate verbal reports on actual or attempted movements (refs. 4, 11, and 16 to 19).

Allegations of gross inaccuracy have also been directed at proprioceptors in skeletal muscles and joints. For example, Crawford (ref. 20) concluded that proprioception is too inaccurate for use in positioning the head. In his test conditions, subjects were required to turn their heads to angles that were named, but not displayed as target stimuli; there was a mean error of about 10° in these conditions. These data are irrelevant to the issue of whether proprioception is used in control. The findings merely show that there are no accurate associations between the names of angles and the specific state of the proprioceptive discharge when the head is held at those angles. This is scarcely surprising, because it is difficult to conceive of situations where such learning could possibly occur.

A previous study (ref. 14) determined the accuracy of pointing an outstretched, hidden arm in the direction of gaze, in the horizontal plane. The mean error and standard deviation were 0.20° and 1.4° , respectively, and the results also indicated that accuracy varied with the frequency of using various coordinated actions. In the present study of vertical forearm movements, the standard deviation of error was 3.51° , centered about a point 1.8° short of the target. The considerable increase of error as compared with the previous experiment may stem in part from the relative infrequency of aligning the forearms, as compared with pointing the arm. Goldscheider (ref. 21) showed that the proprioceptors in proximal joints are more sensitive than those in distal joints. Both findings could help to account for the relatively low accuracy of movements of the forearm, as compared with the whole arm. However, a major function of arm movements is to bring the hand or a held object into the zone of clear, foveal vision, and the accuracy of forearm movement is adequate for the purpose.

In the absence of a visible target, the accuracy of eye and-limb movements is compatible with feedback theory although the level of accuracy does not, as is often supposed, provide a basis for deciding between theories. Accuracy or inaccuracy could equally well be attributed to motor volleys or to feedback impulses.

OUTFLOW THEORY.—According to outflow theory, learned patterns of motor impulses are the only factors effective in control. Limitations of practice time and the storage capacity of the brain preclude the possibility that an indefinite number of patterns could be established for each discriminable movement of each body member. At most, there would be a limited repertoire of stereotyped responses. **This** type of open-chain control cannot be reconciled with extreme variability in primary movements, the extent and direction of residual errors, and the number of subsequent adjustments.

In both experiment 1 and 2 there was considerable variation in the direction and extent of residual errors in primary movements. Subsequent adjustments were often numerous and complex, but tables 2 and 3 show that they reduced the initial error in the vast majority of cases.

The results provide clear evidence for an error-sensing device which is used to correct highly variable **errors**. Outflow theory cannot remotely account for the results, and it is decisively rejected.

THE EFFECTS OF STRESS.—In experiment 1, alcohol consumption that produced a breathalyzer reading of 0.1 percent had a significant ($p < 0.01$) effect of the accuracy of eye-movements. The large reduction in the ratio of correct to incorrect secondary adjustments in recall movements, probably reflected a sharp rise in proprioceptive thresholds.

The subjects in experiment 2 showed surprisingly high resistance to the effects of stress. At the time of taking run 8, the subjects had been without sleep for about 30 hours, but there was no loss of accuracy in aligning the forearms in run 8 as compared with the beginning of practice. The subjects took longer, however, to make adjustments in run 8 than in run 1. There was a further loss of about 3 hours sleep between runs 8 and 9 but the main change was that alcohol consumption produced a breathalyzer level of about 0.06 percent in the latter run. The addition of alcohol stress did not produce a deterioration in performance. Apparently, the alcohol compensated for the extra 3 hours lack of sleep because there was no deterioration in performance between runs 8 and 9. Stresses from several sources are often additive but they may also cancel, as in the interaction of loud noise and sleep deprivation.

INFLOW THEORY

The duration of kinaesthetic reaction time, about 0.2 second in the simple case, has seemed a crucial objection to many writers in precluding continuous proprioceptive control of rapid movements (refs. 3 and 22 to 26). It has seemed impossible to explain how a sensory indication of limb position could be transmitted to high centers and return in time to monitor and terminate movements of shorter duration than kinaesthetic reaction time. The rapid finger strokes of musicians apparently preclude control by a modality in which reaction time plus movement time would limit movement rates to about two per second.

The latency would not preclude parametric feedback, however; proprioception could have a delayed effect, following a movement of incorrect extent, in altering the amplitude of successive movements. Similarly, latency would not rule out the use of proprioception for fine, secondary adjustments.

The problem of latency is largely artificial, stemming in part from the simple, positional control system that is usually postulated as a model. Latency presents no problem if an integral-error control system is used as a model. The rapid component of the proprioceptive discharge is related to the rate of movement. The duration of primary movements is relatively constant at about 0.25 second. The extent of movement could be determined by generating and monitoring an average speed of movement that is appropriate for the distance to be travelled. There would be two main sources of error in controlling primary movements, that is, estimates of the duration of movement and the speed of movement, as defined by the proprioceptive discharge. There is only one main source of error in estimating position from the proprioceptive discharge when making small, secondary adjustments. A clear prediction from the hypothesis of integral error control is that the vast majority of secondary adjustments will reduce the residual errors in primary movements. The present results are in complete agreement with the hypothesis of integral error control, and the model is also compatible with relevant physiological data.

The latency problem has another artificial basis in the persistent confusion between different classes of activity. For example, Lashley (ref. 24) discounted inflow theory because of the apparent impossibility of reconciling kinaesthetic reaction time of about 0.2 second with the rapid finger strokes of musicians at about 12 per second. Reaction times are measured in conditions where the timing and nature of an appropriate response cannot be predicted. In contrast, musicians learn to read ahead, and they memorize sequences of notes so that responses can be prepared before they are needed and reaction delays do not occur.

THE USE OF MODELS.—Ruch (ref. 1) questions the value of mechanical analogies or models of biological processes when the former are devoid of mathematical treatment, but, in fact, mathematical description does not necessarily add anything of value to understanding of the functions and mechanisms of organisms. The major use of models still lies in the primary task of science, which is to classify phenomena.

Many besides Lashley have confused different classes of activity, and models can greatly reduce such confusion. When appropriate responses and their timing cannot be predicted, information theory often provides an appropriate model. A suitable model for the highly practiced activity of musicians is a phonograph. As a result of practice, the appropriate model may change. For example, Hick (ref. 27) found, like others before him, that choice reaction time lengthened with increased number of choices. Hick's unique and invaluable contribution lay in developing an elegant model of the mechanism of decision from the information theory approach. A contrasting approach is provided by Mowbray and Rhoades (ref. 28), who showed that a subject who was very highly trained on a few choices did not show the usual lengthening of reaction time with increased number of choices. They deprecated the use of models on the unfounded and irrelevant grounds that machines, unlike organisms, cannot exhibit adaptability. It is important to note that in discarding relevant models, Mowbray and Rhoades were immediately limited to merely naming phenomena rather than explaining function and mechanism. They "explained" differences in results by contrasting the "lack of practice" of the subject in Hick's study with the "overlearning" of the subject in their experiment. The real problem is of course, to explain the reorganization that must underlie improvements in control, rather than to bury the crucial issue in the vast necropolis of learning theory. Failing agreement on the functions of proprioception there must necessarily be disagreement on the elementary issue of whether patterns of motor impulses, or of sensory stimuli, are the basic content of learning. It has been shown (ref. 12), that in step-input tracking, reaction time lengthens considerably with increased uncertainty of the required direction of response but little, if at all, with an increase in the number of movements of different extent that may be demanded. It can scarcely be supposed that subjects had "overlearned" the control of extent but not of direction; the real difference is best illustrated by the use of models.

The problem of movement control is important in its own right and because a solution would provide the essential basis for understanding related aspects of behavior. Related problems include hand dominance, motor generalization, psychological refractoriness, mirror reversal and transfer effects, and the basic problem of learning, that is, "what is learned?"

Twenty additional years of research have produced no discernable advance on Craik's model of an intermittent correction servo-mechanism for tracking behavior, but he had not solved the general problem of proprioceptive function before his untimely death. It now seems probable that there is continuous monitoring of the speed and direction of movement by negative feedback from the proprioceptors; intermittent corrections are usually based on visual cues. Proprioception probably determines the extent of discrete movements by the use of integral error control and provides the basis for fine, secondary adjustments that do not demand visual attention. In serial movements, proprioception may well provide parametric feedback.

REFERENCES

1. Ruch, T. C.: Motor System. Handbook of Experimental Psychology. S. S. Stevens, ed., John Wiley and Sons, Inc.
2. Craik, K. J. W.: Theory of the Human Operator in Control Systems. I: The Operator as an Engineering System. Brit. J. Psychol., vol. 2, no. 38, 1947, pp. 56-61.

3. Craik, K. J. W.: Theory of the Human Operator in Control Systems. 11: Man as an Element in a Control System. *Brit. J. Psychol.*, vol. 3, no. 38, 1948, pp. 142-148.
4. Helmholtz, H.: Treatise on Physiological Optics. (1867.) Translation by J. P. C. Southall, Ithaca. *Optical. Soc. Am.*, vol. 3, 1925, pp. 243-246.
5. Bell, C.: Idea of a New Anatomy of the Brain. Privately published in 1811. In *Readings of the History of Psychology*, W. Dennis, ed., Appleton-Century-Crofts, 1948, pp. 113-124.
6. Magendie, F.: Experiments on the Functions of the Roots of the Spinal Nerves. (1822.) Reprinted in Olmstead, J. M. D. Francois Magendie. Henry Schumann (New York), 1944.
7. Sherrington, C. S.: Observations on the Sensual Role of Proprioceptive Nerve Supply of Extrinsic Ocular Muscles. *Brain*, vol. 41, 1918, pp. 332-343.
8. Gibbs, C. B.: The Continuous Regulation of Skilled Response by Kinaesthetic Feedback. *Brit. J. Psychol.*, vol. 45, 1954, pp. 24-39.
9. Vince, M. A.: Corrective Movements in a Pursuit Task. *Quart. J. Exp. Psychol.*, vol. 2, no. 1, 1948, pp. 85-103.
10. Woodworth, R. S.: The Accuracy of Voluntary Movement. *Psychol. Monogr.*, vol. 3, no. 2, 1899.
11. Ludvigh, E.: Control of Ocular Movements and Visual Interpretation of Environment. *A. M. A. Arch. Ophthal.*, vol. 48, 1952, pp. 442-448.
12. Gibbs, C. B.: Probability Learning in Step-Input Tracking. *Brit. J. Psychol.*, vol. 56, nos. 2 and 3, 1965, pp. 233-242.
13. Mountcastle, V. B.; Poggio, G. F.; and Werner, G.: The Relation of Thalamic Cell Response to Peripheral Stimuli Varied Over an Intensive Continuum. *J. Neurophysiol.*, vol. XXVI, no. 5, 1963, pp. 807-834.
14. Gibbs, C. B.; and Logan, O.: Tests of the Functions of Proprioception and Interaction of Senses. *Percept. Mot. Skills*, vol. 20, 1965, pp. 433-442.
15. Ditchburn, R. W.: Eye Movements and Visual Perception. *Research*, vol. 12, no. 9, 1956, pp. 466-471.
16. Irvine, S. R.; and Ludvigh, E. J.: Is Ocular Proprioceptive Sense Concerned With Vision? *A. M. A. Arch. Ophthal.*, vol. 15, 1936, pp. 1037-1049.
17. Hammond, P. H.; Merton, P. A.; and Sutton, G. G.: Nervous Gradation of Muscular Contraction. *Physiology of Voluntary Muscle*. *Brit. Med. Bull.*, vol. 3, no. 12, 1956, pp. 214-218.
18. Fender, D. H.; and Nye, P. W.: An Investigation of the Mechanisms of Eye Movement Control. *Kybernetik*, vol. 2, no. 1, 1961, pp. 81-88.
19. Merton, P. A.: The Accuracy of Directing the Eyes and the Hand in the Dark. *J. Physiol.*, vol. 156, 1961, pp. 555-577.
20. Crawford, W. A.: The Perception of Moving Objects. VII: Some Observations on the Presence of Proprioceptive Information From Extraocular Eye Muscles. Memo 150F, Flying Personnel Research Cttee., Air Ministry, 1960.
21. Goldscheider, A.: Untersuchungen uber den Muskelsinn. *Arch. Ant. Psychol.*, 1889, pp. 369-502.
22. Peters, W.; and Wenbourne, A. A.: The Time Pattern of Voluntary Movements. *Brit. J. Psychol.*, vol. 27, 1936, pp. 60-73.
23. Taylor, F. V.; and Birmingham, H. P.: Studies of Tracking Behaviour. 11: The Acceleration Pattern of Quick Manual Corrective Responses. *J. Exp. Psychol.*, vol. 38, 1948, pp. 783-795.
24. Lashley, K. S.: The Problem of Serial Order in Behaviour. In *Cerebral Mechanisms of Behaviour*. Ed. L. A. Jeffress, Soc. Exp. Biol. Symp. No. IV, Cambridge Univ. Press, 1951.

25. Begbie, G. H.: Accuracy of Aiming in Linear Hand Movements. Quart. J. Exp. Psychol., vol. 2, no. XI, 1959, pp. 65-75.
26. Marks, H. J.: Elementary Thinking and the Classification of Behaviour. Science, vol. 3498, no. 135, 1962, pp. 75-87.
27. Hick, W. E.: On the Rate of Gain of Information. Quart. J. Exp. Psychol., vol. 4, 1952, pp. 11-26.
28. Mowbray, G. H.; and Rhoades, M. V.: On the Reduction of Choice Reaction Time with Practice. Quart. J. Exp. Psychol., vol. 1, no. XI, 1959, pp. 16-23.

22. A Neuromuscular Actuation System Model"

D. T. McRuer, R. E. Magdaleno, and G. P. Moore
Systems Technology, Inc.

N68 - 15923

Recently both high-quality physiological data and human-operator describing-function data of low variability and large dynamic range have become available. These data lead to control engineering descriptions for neuromuscular actuation systems which are compatible with the available data and which provide insight into the overall human control structure (e.g., the types of feedback systems used for various inputs). In this paper, some of these physiological and human-operator data are briefly reviewed, and a simple neuromuscular actuation system model is presented.

The physiological data of interest include recent anatomical and physiological data for the muscle spindle and input-output studies of the muscle. These data indicate that simple linear models can describe the basic behavior of these two elements in tracking tasks. Developed further here is the variation in system parameters as a function of average muscle tension and the role of the muscle spindle both as an equalization element and in its effects on muscle tone.

The pertinent human-operator describing-function data include the covariation of high- and low-frequency phase data and the describing-function variation of high-frequency phase with tension.

The simplest neuromuscular model suggested by and compatible with these data is one in which muscle spindles provide both a feedback function, an operating point or bias adjustment, and at least one command path.

The neuromuscular system is a composite of neural and muscular components situated in the spinal cord and the periphery—typically a limb and its neural connections—operating on commands sent from higher centers. We are interested in engineering descriptions of such systems in three respects:

Manual Control Engineering—The basic dynamics of the human operator and the precision of manual control are critically limited by the properties of the neuromuscular system. An understanding of this system has important practical ramifications in determining the effects of control system nonlinearities and sensitivities on manual control.

Control Theoretic—The neuromuscular system is an archetypical adaptive actuation system which, if understood operationally, might serve as the inspiration for analogous inanimate systems with similarly useful properties.

Physiological System Description—The study of the neuromuscular system as a biological servomechanism provides a framework for the interpretation and elaboration of neurophysiological data.

%This paper includes research efforts supported by the Ames Research Center of the National Aeronautics and Space Administration under Contract NAS 2-2824 and by the Air Force Flight Dynamics Laboratory, Research and Technology Division, under Contract AF 33(657)-10835.

In this paper we first summarize some experimental data which have motivated our present model of the neuromuscular system. Then the components of the neuromuscular system are described, and simplified mathematical models for them are developed. These components are connected into a system structure, and its operations for typical control situations are explored. Finally, the ability of the model to account for the experimental data is discussed.

SYMBOLS

| | |
|-----------------|--|
| a | constant in Hill's equation; spindle lead/lag ratio |
| B_c | manipulator damping |
| $B_M(P_0)$ | equivalent damper for muscle system |
| B_N | small signal equivalent damper (under γ_d control) |
| b | constant in Hill's equation |
| C_d | constant of proportionality |
| C_f | constant of proportionality |
| C_{f_s} | constant of proportionality for primary ending firing rate |
| d_f | fiber diameter |
| F | contractive force |
| f | average nerve firing rate in pulses/sec |
| f_0 | operating point average firing rate |
| f_{\max} | muscle firing rate producing tetanus |
| Af_α | incremental alpha motoneuron firing rate |
| $G_M(\gamma_0)$ | muscle and manipulator dynamics |
| i | summation index |
| $j\omega$ | imaginary part of the complex variable, $s = \sigma \pm j\omega$ |
| K_c | controlled element gain; manipulator spring gradient |
| K_{dc} | the dc gain of the open-loop muscle-spindle system |
| K_F | small signal equivalent spring (under γ_s control) |
| K_M | equivalent spring gradient for muscle system |

| | |
|---|---|
| K_N | small signal equivalent spring (under v_d control) |
| K_P | human pilot gain |
| K_S | small signal equivalent spring for the nuclear bag region of a muscle spindle |
| K_{sp} | spindle gain |
| K_α | muscle sensitivity to alpha motorneuron firing |
| L | actual length of muscle |
| L_0 | operating point length of muscle |
| ΔL | length of muscle shortening |
| M | limb + manipulator inertia |
| m | integer |
| P | tension |
| P_0 | operating point tension |
| P_T | tetanic tension |
| P_1, P_2 | particular operating point tensions |
| s | complex variable, $s = \sigma + j\omega$; Laplace transform variable |
| T | time constant |
| T_I | general lag time constant of human pilot describing function |
| T_K, T'_K | lead and lag time constants in precision model of human pilot describing function |
| $\frac{1}{T'_{K_{P_1}}}, \frac{1}{T'_{K_{P_2}}}$ | closed-loop pole for tension P_1 or P_2 |
| $\frac{1}{aT_K}$ | open-loop spindle model pole |
| $\frac{1}{(aT_K)'}$ | closed-loop spindle model pole |
| $\left. \begin{matrix} T_{lag_i} \\ T_{lead_i} \end{matrix} \right\}$ | general lag and lead time constants |
| T_L | general lead time constant of human pilot describing function |

| | |
|--|---|
| $\frac{1}{T_{M_1 P_1}}, \frac{1}{T_{M_1 P_2}}$ | low-frequency muscle-manipulator root for tension P_1 or P_2 |
| $\frac{1}{T_{M_2 P_1}}, \frac{1}{T_{M_2 P_2}}$ | high-frequency muscle-manipulator root for tension P_1 or P_2 |
| T_N | first-order lag time constant approximation of the neuromuscular system |
| T_{N_1} | first-order lag time constant of the neuromuscular system |
| V | velocity of contraction |
| V_0 | operating point velocity |
| ΔV | differential velocity of contraction |
| $Y_c(j\omega)$ | controlled element (machine and display) transfer function |
| Y_p | pilot describing function |
| α | low frequency phase approximation parameter |
| α_C | alpha motorneuron command |
| α_1 | low-frequency phase approximation parameter for tension P_1 |
| α_2 | low-frequency phase approximation parameter for tension P_2 |
| γ | gamma command following gamma motorneuron delay |
| γ_b | gamma bias input due to y_d |
| γ_{b_0} | steady-state value of γ_b |
| γ_c | gamma command input due to y_s |
| γ_{c_0} | steady-state value of γ_c |
| γ_d | gamma input due to dynamic gamma motorneuron |
| γ_0 | total gamma bias |
| γ_s | gamma input due to static gamma motorneuron |
| ζ_M | damping ratio of limb/manipulator system |
| ζ_{M_0} | minimum damping ratio of limb/manipulator system |

| | |
|---------------------------|---|
| ζ_N | damping ratio of second-order component of the neuromuscular system |
| θ | limb rotation |
| σ | real axis of complex plane |
| τ | pure time delay |
| τ_e | effective time delay |
| τ_{e_1} | effective time delay for tension P_1 |
| τ_{e_2} | effective time delay for tension P_2 |
| τ_α | net time delay in the alpha motoneuron pathway |
| τ_γ | net time delay in the gamma motoneuron pathways |
| $\Delta\phi_{\text{low}}$ | incremental low-frequency phase angle |
| ω | angular frequency, rad/sec |
| ω_i | forcing function bandwidth |
| ω_M | undamped natural frequency of limb/manipulator system, $\sqrt{K/M}$ |
| ω_N | undamped natural frequency of second-order part of the neuromuscular system |

Mathematical conventions :

| | |
|--------------|------------------------|
| \doteq | approximately equal to |
| \angle | angle of |
| $ $ | magnitude |
| \uparrow | increase |
| \downarrow | decrease |

For a power quantity (e.g., spectrum) decibels (dB) are expressed as $10 \log_{10}(\)$; for an amplitude quantity (e.g., γ_p) decibels are expressed as $20 \log_{10}(\)$.

HUMAN OPERATOR DATA INDICATING NEUROMUSCULAR SYSTEM EFFECTS

There is a large body of describing function data available for the overall human operator (e.g., refs. 1 to 4) as well as for the neuromuscular system (e.g., refs. 5 to 7). In this paper we are interested in data that relate the effects of average muscle tension or tone on the

neuromuscular-system describing-function data. These effects occur at very high and very low frequencies, and so with measurements limited to midfrequencies they show up primarily in the phase data.

HIGH- AND LOW-FREQUENCY PHASE CONVARIATION.—To describe some recent operator dynamics data, several levels of approximation have been used (refs. 2, 4, and 7). Two of these are shown in figure 1. These "describing function models" are descriptive of the

Precision Model

$$Y_p = K_p \left(\frac{T_L j\omega + 1}{T_T j\omega + 1} \right) e^{-j\omega\tau} \left\{ \underbrace{\left(\frac{T_K j\omega + 1}{T_K' j\omega + 1} \right)}_{e^{-j\alpha/\omega}} \underbrace{\left(\frac{1}{(T_N j\omega + 1) \left(\frac{j\omega}{\omega_N} \right)^2 + \frac{2\xi_N j\omega}{\omega_N} + 1} \right)}_{(T_N j\omega + 1)^{-1} \text{ or } e^{-j\omega T_N}} \right\}$$

where $\alpha \triangleq \frac{1}{T_K} - \frac{1}{T_K'}$ where $T_N \triangleq T_{N1} + \frac{2\xi_N}{\omega_N}$

Approximate Model

$$Y_p = K_p \left(\frac{T_L j\omega + 1}{T_T j\omega + 1} \right) e^{-j\omega\tau} \frac{e^{-j\alpha/\omega}}{T_N j\omega + 1}$$

$$\triangleq K_p \left(\frac{T_L j\omega + 1}{T_T j\omega + 1} \right) e^{-j[\omega(\tau + T_N) + (\alpha/\omega)]} ; \quad \tau_e = \tau + T_N$$

Figure 1.—Describing function models.

approximate model, the lower frequency effects of these third-order terms are approximated by either a first-order lag (so-called neuromuscular lag) or a pure time delay. The latter can be summed with the basic latencies to give an overall time delay τ_e .

The very low frequency characteristics appear in the data primarily as a phase lag. In the precision model these characteristics are represented by the lead divided by lag, $(T_K j\omega + 1)/(T_K' j\omega + 1)$, which is a minimum form suitable to characterize both amplitude ratio and phase data completely for the limited data of extremely high precision (i.e., those data for controlled element forms K_C , K_C/s , and $K_C/(s-1/T)$). For many systems these low-frequency effects can be further approximated, as derived below, by the single-parameter form $e^{-j\alpha/\omega}$. If the low-frequency effects are modeled by transfer characteristics containing m lags and leads, the incremental phase shift due to these will be

$$\Delta\phi_{\text{low}} = \sum_{i=1}^m \tan^{-1} \left(\omega T_{\text{lead}_i} \right) - \sum_{i=1}^m \tan^{-1} \left(\omega T_{\text{lag}_i} \right) \quad (1)$$

At frequencies well above the break frequencies of these lags and leads, that is, $\omega \gg 1/T_{\text{lead}_i}$, $1/T_{\text{lag}_i}$, this can be approximated by

¹The describing function is written in terms of the frequency operator, $j\omega$, to emphasize that this kind of describing function is valid only in the frequency domain and exists only under essentially stationary conditions. For instance, it does not describe the system response to a discrete input, such as a step input (although some of the terms in the model may be compatible with portions of the step-response time history).

$$\begin{aligned}
\Delta\varphi_{\text{low}} &\doteq \sum_{i=1}^m \left(\frac{\pi}{2} - \frac{i}{\omega T_{\text{lead}}} \right)_i - \sum_{i=1}^m \left(\frac{\pi}{2} - \frac{1}{\omega T_{\text{lag}}} \right)_i \\
&\doteq \underbrace{-\frac{1}{\omega} \sum_{i=1}^m \left(\frac{1}{T_{\text{lead}}} - \frac{1}{T_{\text{lag}}} \right)_i}_{\alpha}
\end{aligned} \tag{2}$$

We will use this $e^{-j\alpha/\omega}$ phase characteristic to approximate the midfrequency effects of very low frequency leads and lags.

Our emphasis here is on the α and equivalent time-delay quantities contained in the approximate models exponential phase descriptor term, $e^{-j(\omega\tau_e + \alpha/\omega)}$. Figure 2 illustrates the nature of typical pilot describing-function data and the application of the previously given precision and approximate model forms as descriptors of these data. The typical data shown are from a so-called subcritical task involving the control of a controlled element Y_c consisting of a first-order divergence. The α and τ_e aspects do not affect the amplitude ratio at all, although they are clearly shown in the phase. The $\omega\tau_e$ phase due to time delay dominates the high frequencies, whereas the α/ω phase lag is the major low-frequency effect. Their joint action tends to make the phase look like an umbrella, with α controlling the left side and τ_e the right side; that is, changes in τ_e shift the right side of the umbrella, while changes in α shift the left. Simultaneous increases in both α and $1/\tau_e$ shift the umbrella to the right, whereas decreases shift it to the left.

Some idea of the variation of α and τ_e and their connections is provided in reference 7. Figure 3, which is taken from reference 7, indicates that α and $1/\tau_e$ vary together for the experiments considered there. In terms of the describing-function phase curve shown in figure 2, both ends of the umbrella are shifted together in an adaptive response to forcing-function bandwidth ω_i changes (for $Y_c = K_c/(s)^2$) or controlled-element divergent time constant T changes (for $Y_c = K_c/s(s - 1/T)$).

HIGH-FREQUENCY PHASE VARIATION WITH TENSION.—The effect of average voluntary muscle tension on the supination-pronation response to mechanical impulse inputs has been investigated in reference 5. The relative amount of muscle tension was inferred from a sphygmomanometer cuff attached around the forearm. This was displayed to the subject who then could readily set the reading to any one of five levels. These experiments were carried out using irregularly spaced mechanical impulses delivered, without warning, by a pendulum. The manipulator restraint consisted of an inertia several times larger than that of the arm, and the subject was asked to resist the perturbing influence of the pendulum-produced disturbance on the load.

The transient response resembled that of a dominant second-order system with light damping. Figure 4, taken from reference 5, shows the upper pole position of a complex pair fitted to the transient response for the five tension values (apparently for only one subject). In general, increasing mean tension increases the natural frequency of these roots but leaves the damping ratio relatively unchanged. We are interested in the effect of this trend on the high-frequency phase which can be found by noting that, as in figure 1,

$$\frac{1}{1 + \left(\frac{2\zeta_N}{\omega_N} \right) j\omega + \left(\frac{j\omega}{\omega_N} \right)^2} = \frac{1}{1 + \left(\frac{2\zeta_N}{\omega_N} \right) j\omega} \quad (3)$$

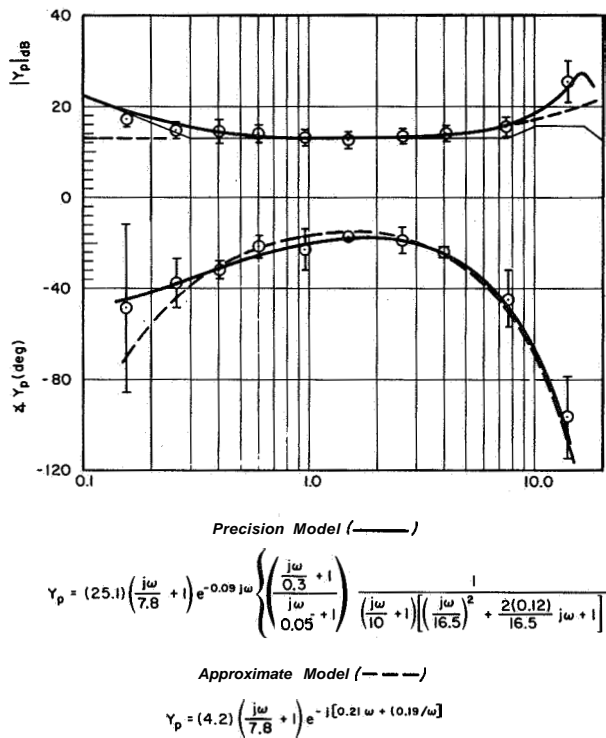


Figure 2.—Typical pilot describing-function data and models ($Y_c = K_c/(s-2)$; $\omega_i = 4.0$ rad/sec).

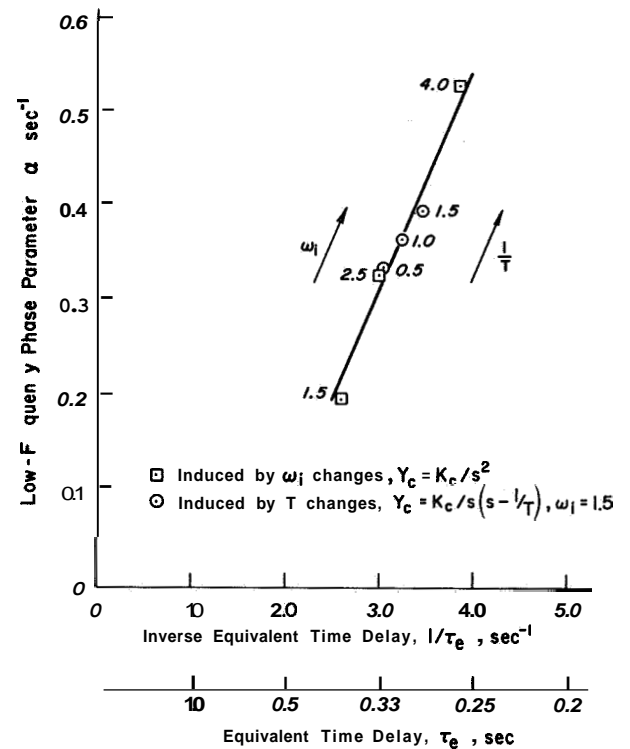


Figure 3.—Connection between equivalent time delay and low-frequency phase lag.

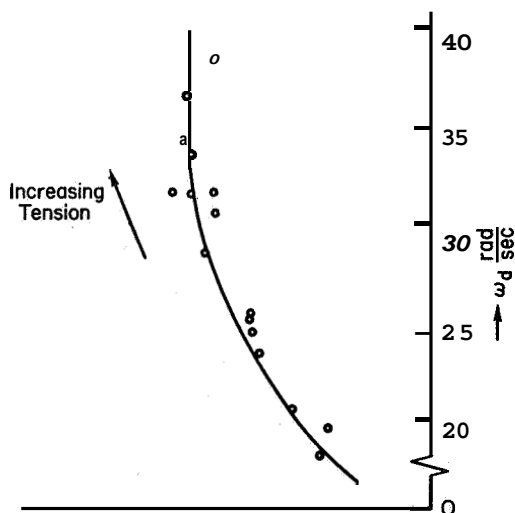


Figure 4.—Root locus of the positive pole of a complex pair fitted to the transient response (ref. 5).

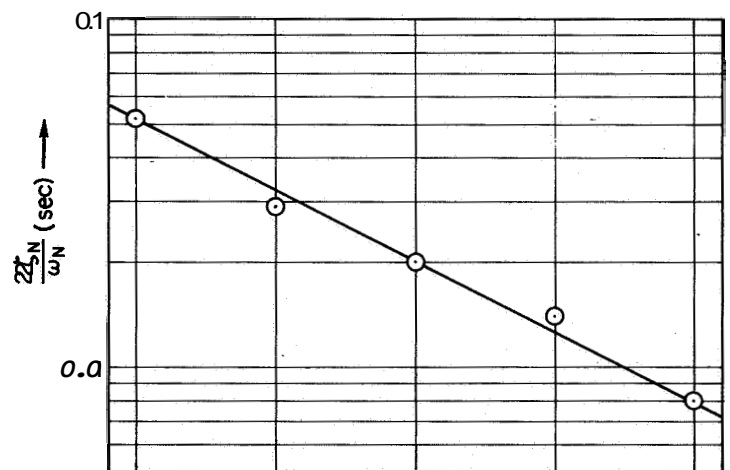


Figure 5.—Effective time constant as a function of inferred tension.

for frequencies below ω_N . Since all values of ω_N are at the extreme upper end of the measurement bandwidth (see fig. 2), then the approximate phase of these complex roots behaves as a simple lag with an effective time constant given by $2\zeta_N/\omega_N$. In the approximate model this is lumped into the overall time delay τ_e .

Reference 5 also contains averaged data for three subjects which we can use to calculate the effective time constant for the five tension values (see fig. 5). The decrease in time constant as tension increases will move the high-frequency phase to the right.

DESCRIPTION OF THE COMPONENTS

To describe the physiology of the neuromuscular system is very difficult because of the enormous complexities of the system and the many unknowns still present despite recent advances (refs. 8 to 10). In spite of these difficulties, we shall attempt here a broad and superficial coverage of some workings of the neuromuscular system elements, expressed in engineering terms. The components that shall be described are, in sequence motor units, sensory units, and interconnection and amplification elements. Only those elements which are believed to be important in the manipulation of spring-restrained low-inertia manipulators with negligible nonlinearities are considered.

MOTOR UNITS.—The smallest functional entity in a motor system is a motor unit, which is illustrated schematically in figure 6. It consists of a motoneuron cell body, located in the ventral horn of the spinal cord; its axon², which is a single-fiber signal-transmission line from the cell body to the muscle; and the group of muscle fibers connected to the terminal branches of the axon. Motor units differ primarily in the number of muscle fibers (innervation ratio) served by one motoneuron. The innervation ratio is a limiting factor in the attainable precision of muscle control. For muscles involved in very precise and finely graded movements, such as those of the eye, the innervation ratio can be as low as about five. In contrast, large muscles which are used primarily in gross movements, such as the biceps, may have hundreds or thousands of muscle fibers per motoneuron. The individual fibers of muscles are grouped into bundles which, in turn, are assembled to make up the complete muscle. There may be as many as several million fibers in a muscle. The fibers of a motor unit are widely dispersed in several bundles, so the action of a single motor unit is spread throughout the muscle.

The muscle responds to the commands carried by the motoneuron. These responses are transmitted along the axon as brief (≈ 1 msec) electrical pulses of essentially invariant shape and duration, which travel at rates of up to 100 m/sec. Arriving at the axon muscle interface ("neuromuscular junction") they trigger the release of a chemical agent which induces the generation of a similar pulse in the muscle fiber; the pulse spreads over the fiber, triggering a sequence of events leading ultimately to contraction of the muscle fiber.

The net result of all this activity is a brief contractile response of the muscle fiber, called a "twitch." A typical twitch time history for isometric conditions (limb constrained so that the muscle system is of essentially constant length) is shown in figure 7. This amounts to the impulse response of the muscle-fiber portion of a motor unit. It is also the weakest possible natural movement involving that muscle. The twitch shown is typical of a fast muscle, with a contraction time of about 40 msec and a far longer decay.

²These axons to skeletal muscle fibers are the largest of the motor fibers and have diameters which are in the "alpha" range (12 to 20 μ). Hence these motoneurons are sometimes called "alpha motoneurons."

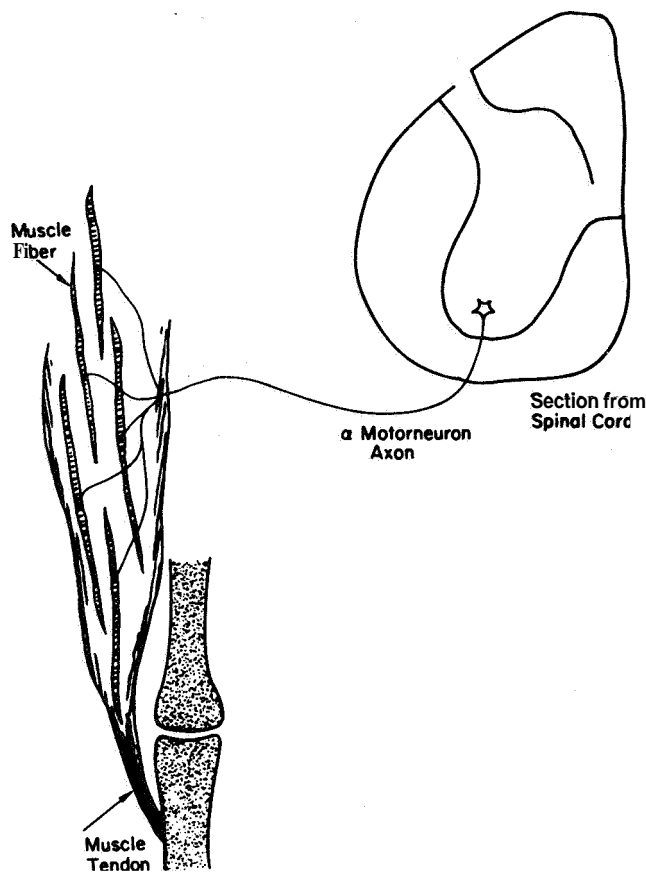


Figure 6.—Diagram of a motor unit.

directly related to the number of active motorneurons and to the rate of spike discharge in these motorneurons. This provides for an enormous dynamic range in a given muscle.

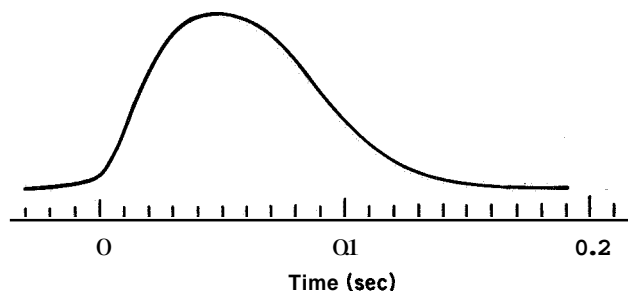


Figure 7.—Impulse response of a muscle fiber.

in equation form for muscles is Hill's equation. It has been found suitable for a wide range of muscle types, including arm muscles in man (ref. 11). Hill's "characteristic equation" (refs. 11 and 12) is

$$(F+a)V=b(P_T-F) \quad (4)$$

In a single motor unit a low repetition rate of action potentials arriving at the motor endplate causes a train of essentially independent twitches. In each of these, the mechanical response lasts far longer than the action potential stimulus. As the frequency of excitation increases, a second action potential will arrive while the mechanical effect of the prior stimulus still persists, causing a mechanical summation or fusion of contractions. Up to a point the degree of summation increases as the stimulus interval becomes shorter, although the summation effect decreases as the interval between the stimuli approaches the refractory period of the muscle. Maximum tension occurs when the excitation frequency is so high that the degree of summation approaches zero. This limiting response is called a tetanus, and the tension developed is about four times that of a single twitch.

In an actual muscle system there are many motor units energized by many coordinated motorneurons. Thus, when more force is needed, not only can a single motor unit be fired more rapidly, but other motor units can be activated and their motorneurons made to discharge more frequently. By this means, the degree of tension in a muscle is

Although deficient in detail, the preceding summary provides a broad qualitative picture of the major electrical and chemical events involved in the translation of a motorneuron command to a muscle response. The picture can be put into analytical form by treating existing data on muscles as if they were force/speed characteristics of an actuator. This can be done using actual data or, alternatively, by using equations fitted to actual data. The second scheme is by far the more convenient. The most popular data summary

where F is the force exerted by the muscle during a contraction with a velocity V , P_T is the isometric ($V=0$) or maximum (tetanic) tension in the muscle, and a and b are constants. Note that P_T depends upon muscle length—this is taken into account in a later paragraph. In this equation FV is the power required to do mechanical work and aV may be considered to be the power dissipated in heat within the muscle.

Hill's equation is often presented in terms of a force/velocity relationship,

$$(F+a)(V+b)=b(P_T+a) \quad (5)$$

In this connection a and b become so-called force and velocity constants. In terms of the modified force and velocity variables, $(F+a)$ and $(V+b)$, the relationship is hyperbolic (see fig. 8), implying constant "power" in these coordinates for a constant isometric tension. Referring to figure 8, Hill's equation has been found to be most accurate for shortening conditions and to depart from the hyperbolic relationship for lengthening velocities. The relationship tends to be more linear and somewhat steeper than an extrapolation from Hill's equation would show. This results in a change in slope in the region of most interest (small positive and negative velocities). However, in the context of an agonist-antagonist muscle pair one muscle is lengthening while the other is shortening. Thus, the overall force/velocity relationship is continuous.

Also shown in figure 8 are curves (ref. 13) for the force/velocity relationship for a number of different levels of average electrical activity measured from surface electrodes.

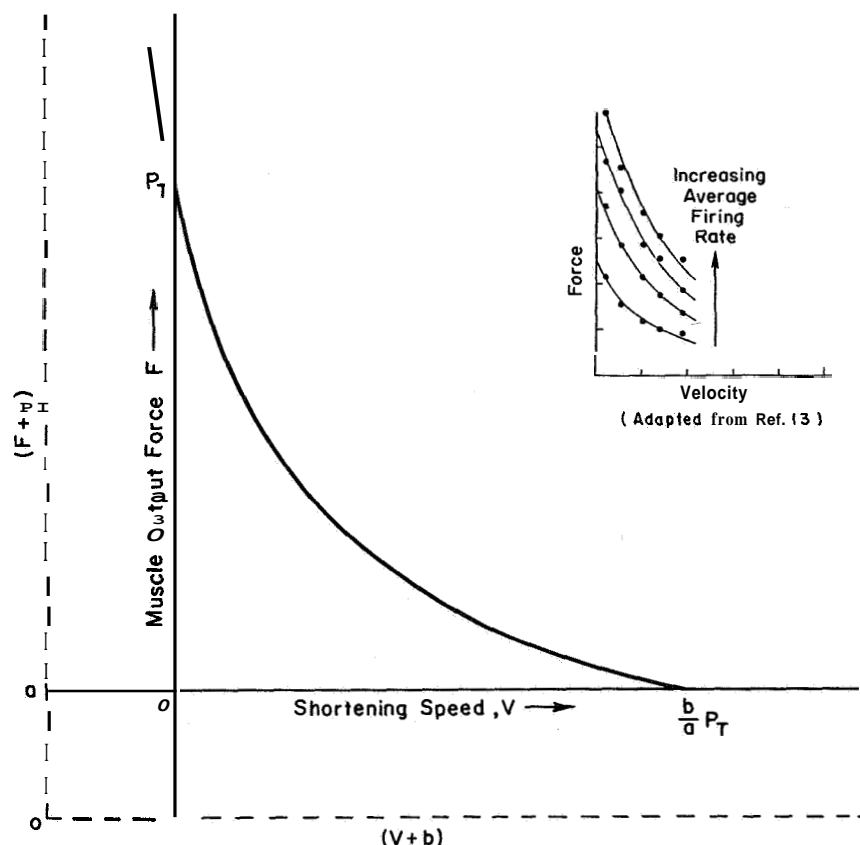


Figure 8.—Force-speed relationship for muscle, $P_T = \text{Constant}$.

These indicate that Hill's equation also applies if P_T is replaced by P , the isometric tension in the muscle due to the average firing rate f of the motor unit ensemble. Note also that the slope of the force/velocity curves increases for increasing average firing rate; that is, the effective damping varies with muscle activity.

When written in the form desired for our purposes, Hill's equation is

$$F = \frac{bP}{b+V} - \frac{aV}{b+V} \quad (6)$$

We are interested primarily in situations where the muscles involved execute small perturbations about the operating point $P=P_0$, $V_0=0$. For these circumstances, the velocity V can be considered small and equation (6) can be expanded in a Taylor series. Assuming that terms of the second-order or higher in perturbation quantities are negligible yields

$$F = F(P_0 + \Delta P, V_0 + \Delta V) \doteq F(P_0, V_0) + \left[\frac{\partial F}{\partial P} \Delta P + \frac{\partial F}{\partial V} \Delta V \right]_{\substack{P=P_0 \\ V=V_0=0}} \quad (7)$$

Evaluating the partial derivatives and setting $\Delta V=V$ yields

$$F \doteq \frac{P_0 + \Delta P}{P} - \left(\frac{a + P_0}{b} \right) V \quad (8)$$

The general tension P in equation (8) is a function of both length and average firing rate as shown in figure 9. The zero and maximum lines on this figure are from reference 14, whereas the dashed lines for other average firing rates represent an interpretation of the data in reference 15. For a control task in which the muscle will have an average tension P_0 , average firing rate f_0 , and average or rest length L_0 , we can describe small deviations about the operating point with a Taylor series expanded about $f=f_0$ and $L=L_0$. Keeping only terms of first order in the perturbation quantities yields (for the tension P near P_0)

$$P = P(f_0 + \Delta f, L_0 + \Delta L) \doteq P_0(f_0, L_0) + \left[\frac{\partial P}{\partial f} \Delta f + \frac{\partial P}{\partial L} \Delta L \right]_{\substack{f=f_0 \\ L=L_0}} \\ \doteq P_0 + C_f \Delta f - \frac{\partial P}{\partial L} \Delta L \quad (9)$$

where ΔL has been defined as positive in the direction of muscle shortening so that $d(\Delta L)/dt$ will equal shortening velocity. Thus the muscle output force contains a steady component, a component sensitive to changes in firing frequency, and an effective spring element. Substitution of equation (9) into equation (8) gives the result

$$F \doteq P_0 + C_f \Delta f - \left(\frac{a + P_0}{b} \right) V - \frac{\partial P}{\partial L} \Delta L \\ \doteq P_0 + C_f \Delta f - B_M V - K_M \Delta L \quad (10)$$

In terms of an analogous physical system the linearized equation for a muscle given previously corresponds to a force source, $P_0 + C_f \Delta f$, coupled to a parallel spring/viscous-damper combination. Such a model is generally sufficient when the muscle is operating near or below the natural muscle length in the body. When the muscle is substantially greater in length, the elastic element may become negative. Except for this consideration, the equivalent elastic element has approximately constant gradient. The damper element, on the other hand, has a damping coefficient which is linearly related to the operating point tension. Consequently, the effective damping is also directly proportional to f_0 , the steady-state firing rate, since $P_0 = C_f f_0$.

Skeletal muscles can only contract actively, so movements involving high-grade skill, such as tracking, generally require coordinated groups of muscles. The simplest of these is an agonist/antagonist pair connected at opposite ends of a first-class lever to provide rotary motion. For rotation to occur one muscle must contract while the other extends. If the opposing muscles each have a steady-state tension in the static situation caused by some steady-state or average firing rate, f_0 , rotation can be accomplished by increasing the firing rate for the contracting muscle by the increment Δf while simultaneously decreasing the firing rate in the antagonist by about the same increment.

The actual muscle system involved in almost any complex limb motions is seldom, if ever, as simple as that described above. However, the same principles hold for each agonist/antagonist pair involved and a grand summation can be made of all the pairs contributing to the actual limb motion of interest.

In tracking actions the muscle system operates in conjunction with a manipulator, which is ordinarily restrained by both linear and nonlinear mechanical elements. If linearity is assumed here, the load dynamics of a wide variety of practical manipulators can be characterized adequately by spring,

manipulator load dynamics for the

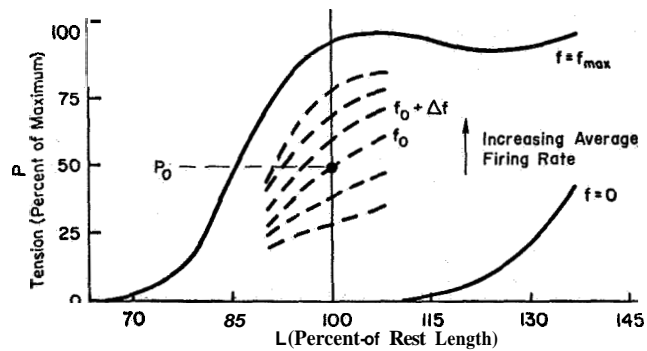
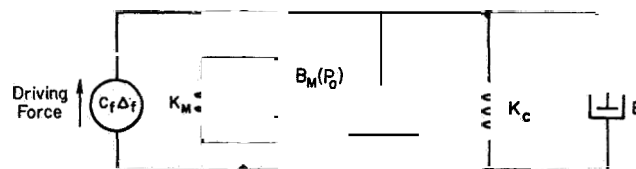


Figure 9.—Presumed tension-length curve (from refs. 14 and 15).

MUSCLE ACTIVE AND PASSIVE CHARACTERISTICS

MANIPULATOR CHARACTERISTICS



$$\frac{\theta}{\Delta f} = \frac{\frac{C_f}{M}}{s^2 + \left[\frac{B_M(P_0) + B_c}{M} \right] s + \frac{K_M + K_c}{M}} \quad (11)$$

For a particular limb/manipulator system the undamped natural frequency

$$\omega_M = \sqrt{\frac{K_M + K_c}{M}} \quad (12)$$

will be constant as long as the motions do not require muscle lengths which depart too far from L_0 . The damping characteristics, on the other hand, will vary directly with the average tension P_0 , which is subject to enormous variation. Consequently, the damping ratio

$$\zeta_M = \frac{B_c + B_M(P_0)}{2 \sqrt{(K_M + K_c)M}} \quad (13)$$

can vary greatly. With a normal subject the minimum average tension in movements having zero mean is $P_0=0$, so the minimum limb/manipulator system damping ratio will be

$$\zeta_0 = \frac{B_c}{2 \sqrt{(K_M + K_c)M}} \quad (14)$$

For precision movements in tracking there is always some average tension acting, and ordinarily this is sufficient to make the damping ratio considerably greater than unity. For this case the appropriate forms for equation (11) are, therefore, either

$$\frac{\theta}{\Delta f} = \frac{\frac{C_f}{(K_M + K_c)}}{\left(\frac{s}{\omega_M}\right)^2 + \frac{2\zeta_M}{\omega_M}s + 1} \quad (15)$$

or, when $\zeta_M > 1$,

$$\frac{\theta}{\Delta f} = \frac{\frac{C_f}{(K_M + K_c)}}{(T_{M_1}s + 1)(T_{M_2}s + 1)} \quad (16)$$

The dynamics of this equivalent system are illustrated for two cases of tension by the $j\omega$ -Bode diagram of figure 11. From this figure it is apparent that the effect of changing the viscous characteristic of the muscle group is to decrease the high-frequency time constant, T_{M_2} , and increase the low-frequency time constant, T_{M_1} . In the process the width of the -20 dB/decade portion on the limb/manipulator system Bode diagram is increased. As will be seen later these changes, due to an increase in steady-state tension, have most important consequences on the neuromuscular system dynamics.

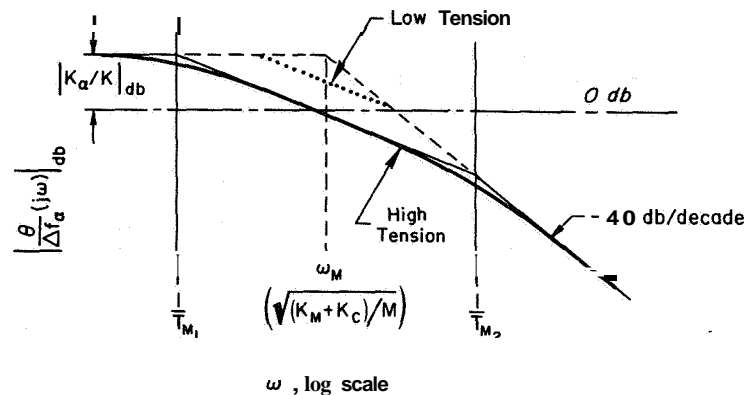


Figure 11. $-j\omega$ -Bode diagram for limb/manipulator dynamics.

MUSCLE SPINDLES

Spindle Anatomy.—Much of the control of neuromuscular behavior in the periphery is dependent on a complex organ located in most muscles of the body, the muscle spindle. It is in itself a complex neuromuscular integrative system receiving a continuous set of motor control and command signals from the central nervous system, and sending a constant stream of sensory information signals via its several paths back to the central nervous system.

A typical muscle may have 50 to 80 of these organs, embedded at various points among the tension-producing ("extrafusal") muscle fibers of the main muscle mass (ref. 16). A typical spindle is elongated in shape, may be several millimeters in length, and has an orientation parallel to that of the extrafusal muscle fibers. They may be arranged in isolation, in tandem with each other, or may be found in conjunction with other specialized receptor structures of the muscle.

Since the influence of the spindle in neuromuscular control is more far-reaching than previously suspected, it is important to review some of its basic anatomical and physiological features. These have been documented in several recent journals and symposia (refs. 17 to 19) and in the following discussion we shall present a summary of this work in which the complexities of spindle structure and function not relevant to the present discussion are either simplified or omitted.

A highly simplified diagrammatic view of a muscle spindle is shown in figure 12. This shows the central axis of the spindle which consists of a globular nuclear bag region connected to either pole of the spindle by means of a pair of nuclear bag fibers, which are themselves typical striated muscle fibers ($\approx 25 \mu\text{m}$ in diameter).³ These nuclear bag fibers are known as "intrafusal fibers," and are never observed to contribute directly to the development of tension in the muscle. Rather, they appear to be motor fibers related solely to control within the spindle itself. From their microscopic appearance the bag fibers appear to be normal striated muscle fibers and hence would be expected to have dynamic and mechanical properties similar to those of the extrafusal fibers discussed previously.

³Chain fibers, which lack the nuclear bag region, have been observed and discussed (ref. 17). They have smaller diameters and are connected in series and in parallel with the nuclear bag fibers. All spindles have at least one nuclear bag fiber, but may not have any chain fibers. Thus we shall assume that the chain fibers are not vital to the spindle's basic behavior.

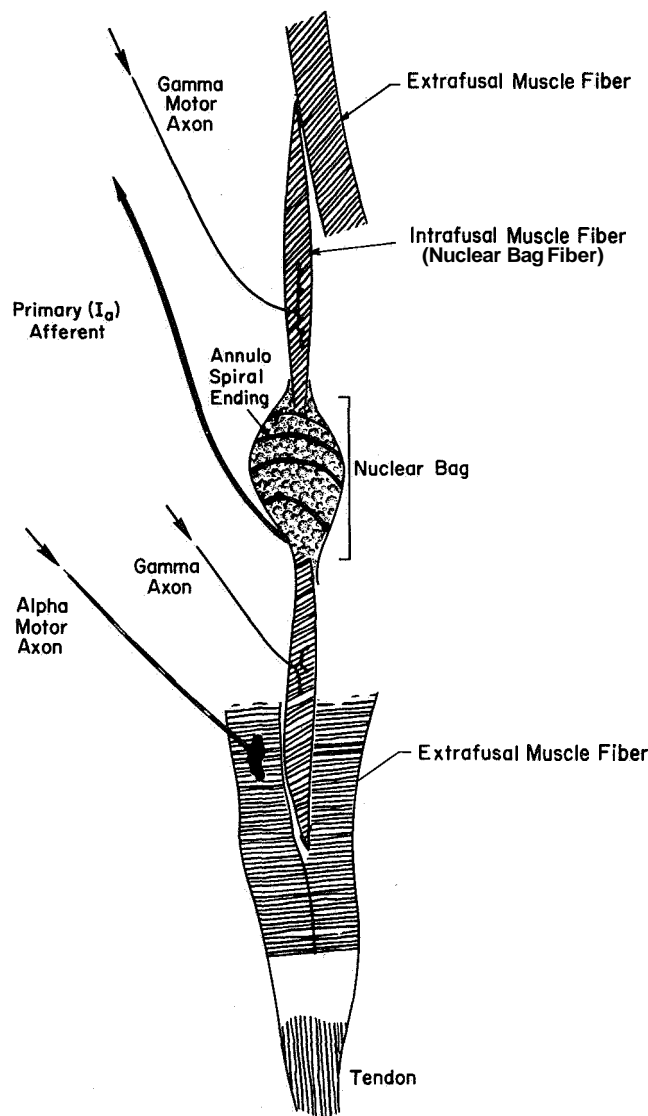


Figure 12.—Idealized nuclear bag muscle spindle.

The nerves supply the intrafusal fibers have their cell bodies situated in the spinal cord and are connected to them by **axons** which have a diameter somewhat smaller than that of the axons connected to extrafusal muscle fibers. To distinguish these two classes of motor cells, axons and muscle fibers, the terms "gamma" (referring to the spindle motor system) and "alpha" (referring to the main muscle motor system) are used.

Of equal importance to these effector or command paths, however, are the sensory fibers which originate at the spindle and which generate pulse trains carrying information back to the spinal cord. In a typical spindle, there will usually arise one large axon whose principal termination winds around the nuclear bag region (a region of the spindle without muscle or contractile elements). These are the primary or annulospiral endings. The type Ia axons serving these have large diameters (12 to 20 μm) and do not have terminations with other spindles. A second type of sensory ending is the flower-spray or secondary ending (not shown in fig. 12). Each axon can terminate in several endings along the intrafusal fiber. These type II axons are about half the diameter of the type Ia afferents. They never appear to terminate in the nuclear region of the bag fiber, but the same type II axon can have terminations in more than one spindle.

In the present paper we are concerned only with the behavior of the larger diameter primary endings—how they respond to changes in muscle length and gamma activity.

Before summarizing the physiological behavior of the primary ending, we must first indicate something about the relation between mechanical events at the nuclear bag region and the electrical events in the sensory annulospiral nerve ending. Much evidence has been accumulated which indicates that mechanical deformations of sensory terminals lead to the development of electrical potential fields at the terminals which are directly proportional to the strength of the deformation (ref. 20). These generator potentials in fact are very accurate mappings of the forces operating on the terminals and can follow rather high frequencies of change in the deforming stimulus. The fields are an inherent property of the receptor membrane itself, and the fields are utilized by the sensory axon in the production of nerve impulses in specialized triggering regions near the receptor endings. Nerve impulses, in fact, are generated at a rate directly proportional to the magnitude of the generator potential; hence there is a continuous transmission of impulses at a frequency which is a linear function of that

potential and hence of the strength of the deformation. The system exhibits a high degree of sensitivity to length changes; a significant shift in firing frequency can result from length changes of only a few microns (ref. 21). We can use this relationship to reconstruct the time course of tension changes at the nuclear bag region from observed trains of nerve impulses.

Physiology of the Spindle Primary Ending.—In this section we briefly consider some of the data concerned with input/output relations in the primary (annulospiral) ending of the spindle. First, we consider the steady-state relations between the firing frequency in the spindle Ia axon as a function of muscle length and how this is influenced by gamma activity. Next we present data regarding the response of the primary ending to transient changes in muscle length and how such responses are modified by gamma activity. Finally, we consider the response of the primary ending to gamma stimulation alone.

On the basis of these data we present a small-signal equivalent lumped-parameter mechanical model of the spindle whose parameters are under gamma control.

The primary ending of a spindle usually shows some spontaneous discharge even when the extrafusal muscle fibers are at their normal resting body length. This is presumably due to a small amount of residual tension in the spindle. Even in the absence of any motor signals from the cord, this spontaneous rate of firing in the spindle will increase monotonically as a function of increasing muscle length (from a few pulses per second to a hundred or more pulses per second). This arises because the disposition of the spindle within the muscle serves to transmit length changes in the muscle to the bag region where the change is reflected as an increase in bag tension. Conversely, shortening of the muscle (either passively or in response to an alpha motor command signal) will reduce the tension on the bag and hence reduce the spindle Ia sensory fiber firing frequency.

Figure 13 shows some typical plots of spindle sensory receptor firing frequency as a function of muscle length. Over a considerable range this relation is linear.

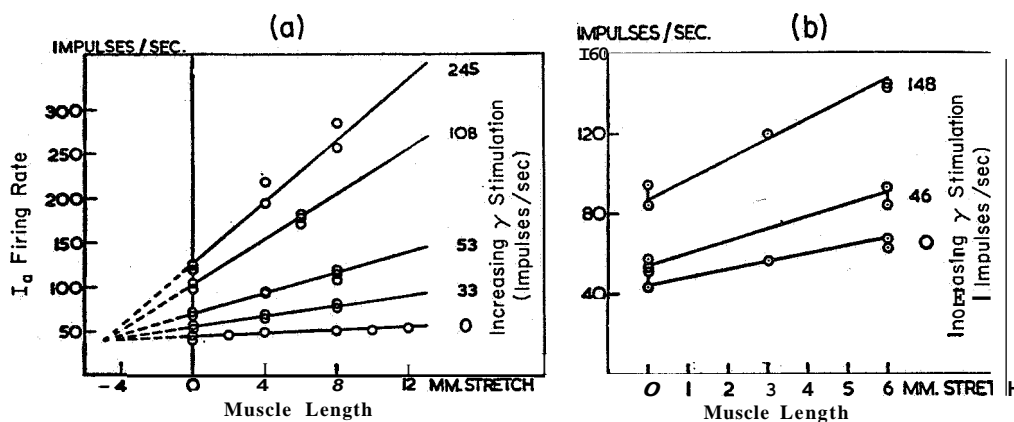


Figure 13.—Relationship between steady-state length and Ia firing rate for various gamma fiber stimulation rates (reproduced from ref. 23).

Recent studies of single gamma fibers ending on spindles whose primary endings were being monitored have shown that repetitive stimulation of certain gamma fibers will produce a response which consists of a general increase in the firing rate of the primary ending (relative to its control rate) as a function of muscle length. This may result in (1) a simple translation of the length-frequency curve upward (fig. 13(a)), or (2) a family of curves of increasing slope whose relations are still essentially linear (fig. 13(b)), or (3) in a combination of both (refs. 22 and 23). Each family of curves is a function of the stimulation frequency of the gamma fiber.

The shift in length-frequency relation is usually unaccompanied by any significant increase in the velocity sensitivity of the fiber (see below); that is, the gamma fiber has influenced only the static gain of the primary ending; such a gamma fiber is designated as a static fusimotor fiber.

Thus, figure 13 shows that the discharge frequency of the primary fiber as a function of muscle length remains linear under static gamma stimulation. The base firing frequency at zero extension increases as gamma firing rate increases; hence the steady-state behavior can be described by an equivalent variable-slope spring element plus a bias level, both under static gamma control.

In addition to this static length-firing frequency relation, the primary ending also shows a strong velocity sensitivity to stretch or release in which its firing frequency, with rapid lengthening, can increase to several hundred pulses per second, regardless of length, and characteristically drops abruptly to zero for many seconds when the muscle is allowed to shorten.

This velocity-sensitive property can also be influenced by gamma stimulation, but the population of gamma fibers with this capacity is distinctly different from that which influences the static relation. Gamma fibers which increase the velocity-sensitive or dynamic part of the response to changes in muscle length with only a small effect on the static gain are thus called dynamic fusimotor fibers. The effect of stimulating these is also dependent on the stimulating frequency. Thus, the primary ending exhibits an inherent relation between firing rate, length, and velocity; the static gamma fiber increases its discharge for a given length (but in fact may decrease its relative sensitivity to stretch). Conversely, the dynamic fiber greatly increases the firing frequency of the primary ending during stretch, but leaves its initial and final frequencies essentially unchanged.

Our interpretation of the present knowledge of the behavior of the primary sensory ending can be summarized by the simplified mechanical network in figure 14. Here we confine our attention to the nuclear bag fiber. The intrafusal fibers are represented as typical skeletal muscle fibers, that is, as an elastic element K_F in series with a parallel elastic-viscous unit, K_N and B_N . The nuclear bag region is represented as a simple elastic element K_S .

Since the steady-state firing rate and the slope of the length-frequency curves are under static gamma control, a static gamma γ_S driven force generator and a variable K_F are indicated conjecturally in the diagram. Similarly, B_N and K_N are indicated as variable under dynamic gamma control, and are in parallel with a dynamic gamma (γ_d) driven force generator. K_S is assumed fixed. Referring to this figure, we see that a step change in muscle length or a step change in static gamma activity will change the primary ending firing rate as in a typical lead/lag network step response (ref. 8). Step changes in dynamic gamma activity may produce some changes in sensory fiber firing rate, but these will show a smooth rise without overshoot. The principal effect, however, will be to alter markedly subsequent responses to static gamma stimulation or muscle length changes. The force generator γ_S (shown dotted) can be replaced by the effective length input γ_c at least as far as the response at X_S (stretch of the nuclear bag ending spring, K_S) is concerned.

A step change in the effective-length input signal (via a step in γ_c or L) will initially be taken up across K_F and K_S , since the damper will not allow sudden position changes. Subsequently, the length change across K_S and K_F will be redistributed between the three springs (K_F , K_N , and K_S), thus reducing the displacement across K_S . The primary ending firing rate then will respond in proportion to deformation of X_S , which will have the form

$$A_f = C \quad X = K_{sp} \left(\frac{T_K s + 1}{a T_K s + 1} \right) (\gamma_c - L) + \frac{C_d \gamma_d}{a T_K s + 1} \quad (17)$$

discharge from the gamma motor system. For example, the steady-state signals in gamma motoneurons which serve both agonist and antagonist muscles via their spindle systems may introduce tension preloads into the opposing muscles, which, as discussed previously, can exert profound changes on the neuromuscular actuation system time constants. In addition, such biases increase the spindle sensitivity by increasing the mean spindle discharge rate per unit deflection and by reducing any rectification effects caused as the spindle tends to become slack.

SIMPLIFIED NEUROMUSCULAR SYSTEM AND CLOSED-LOOP DYNAMICS

SYSTEM DESCRIPTION.—Although the microscopic details of the human operator's actuation system for even the simplest of motions are enormously complicated from a component standpoint, the actions of the overall system can still be modeled simply if component ensembles are used. The connections between these ensembles will depend greatly on the type of neuromuscular system motions involved. In this paper we are concerned primarily with neuromuscular system operations in which the command inputs are random and the motion outputs are exerted on spring-restrained low-inertia manipulators. For this situation an appropriate neuromuscular system can be made up by connecting ensembles of the components described in the previous section into the equivalent system shown in figure 16. For other physical situations different neuromuscular system block diagrams will apply (e.g., see refs. 3 and 10).

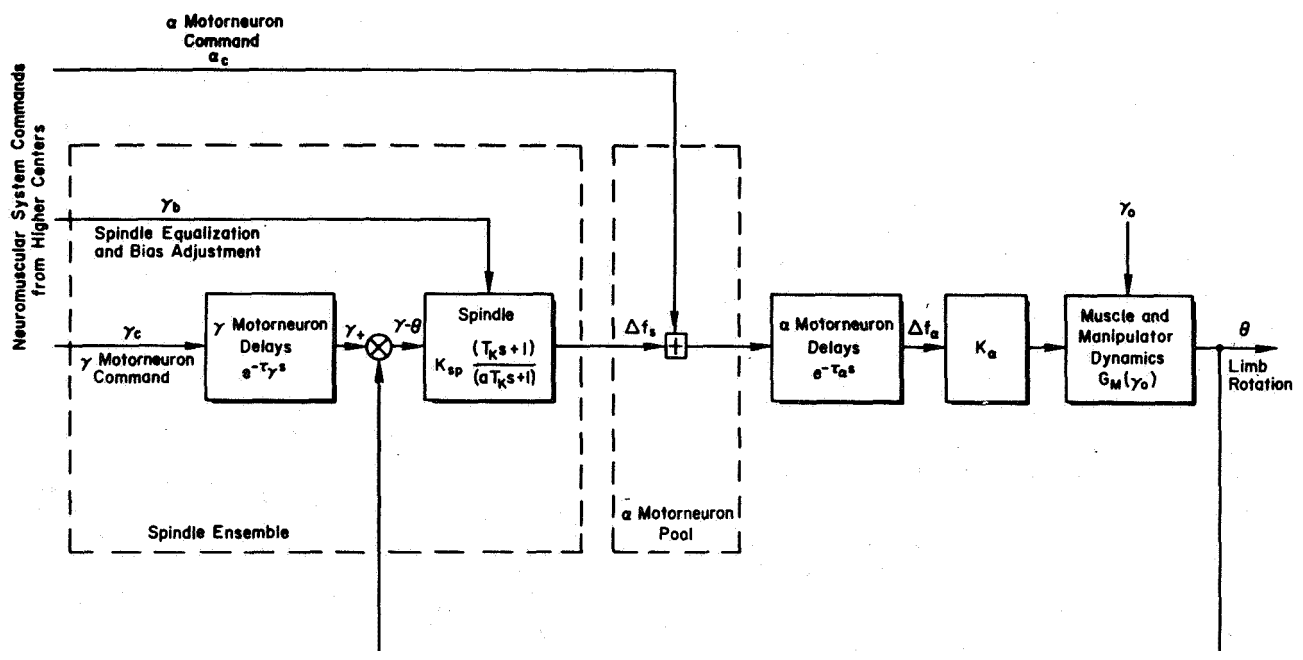


Figure 16.—Elementary neuromuscular system model.

The block diagram of figure 16 shows perturbation operations about steady-state operating points. Consequently, all the signals indicated can be either positive or negative and the agonist/antagonist relationships are subsumed in the composite diagram (a block diagram of an agonist/antagonist operation with absolute level signals can be constructed by simply

duplicating the figure 16 diagram and then connecting the two back to back with the limb rotation θ being a sum of the two outputs).

The spindle ensemble provides in one entity the feedback of θ ; some series equalization, $K_{sp}(T_K s + 1)/(a T_K s + 1)$; the source of one command to the system γ_c ; and a means of spindle equalization and bias adjustment γ_b . The spindle output differential firing rate Δf_s is summed with an alpha motorneuron command α_c with the result, after conduction and synaptic delays, being an incremental alpha motorneuron firing rate Δf_{α} . This in turn operates the muscles and manipulators, giving rise to the limb rotation which is then sensed by the spindle ensemble.

The effective damping in the muscle and manipulator dynamics transfer function, G_M , is an operating point adjustment set by the total gamma bias γ_0 . The bias can derive from either the steady-state values γ_{b0} and γ_{c0} , or both, from which the γ_b and γ_c motorneuron signals represent perturbations. This dependence of muscle manipulator damping on steady-state gamma motorneuron activity is indicated in the block diagram by the γ_0 input into the G_M block. Here we are presuming that the alpha motorneuron command is not involved.

CLOSED-LOOP DYNAMICS—EFFECTS OF AVERAGE TENSION.—The closed-loop dynamic characteristics of the neuromuscular system for γ_c inputs are indicated in figure 17 for two levels of tension. For a given steady-state tension setting, the closed-loop roots are indicated on the conventional root locus and the closed-loop real loops on the Bode root locus

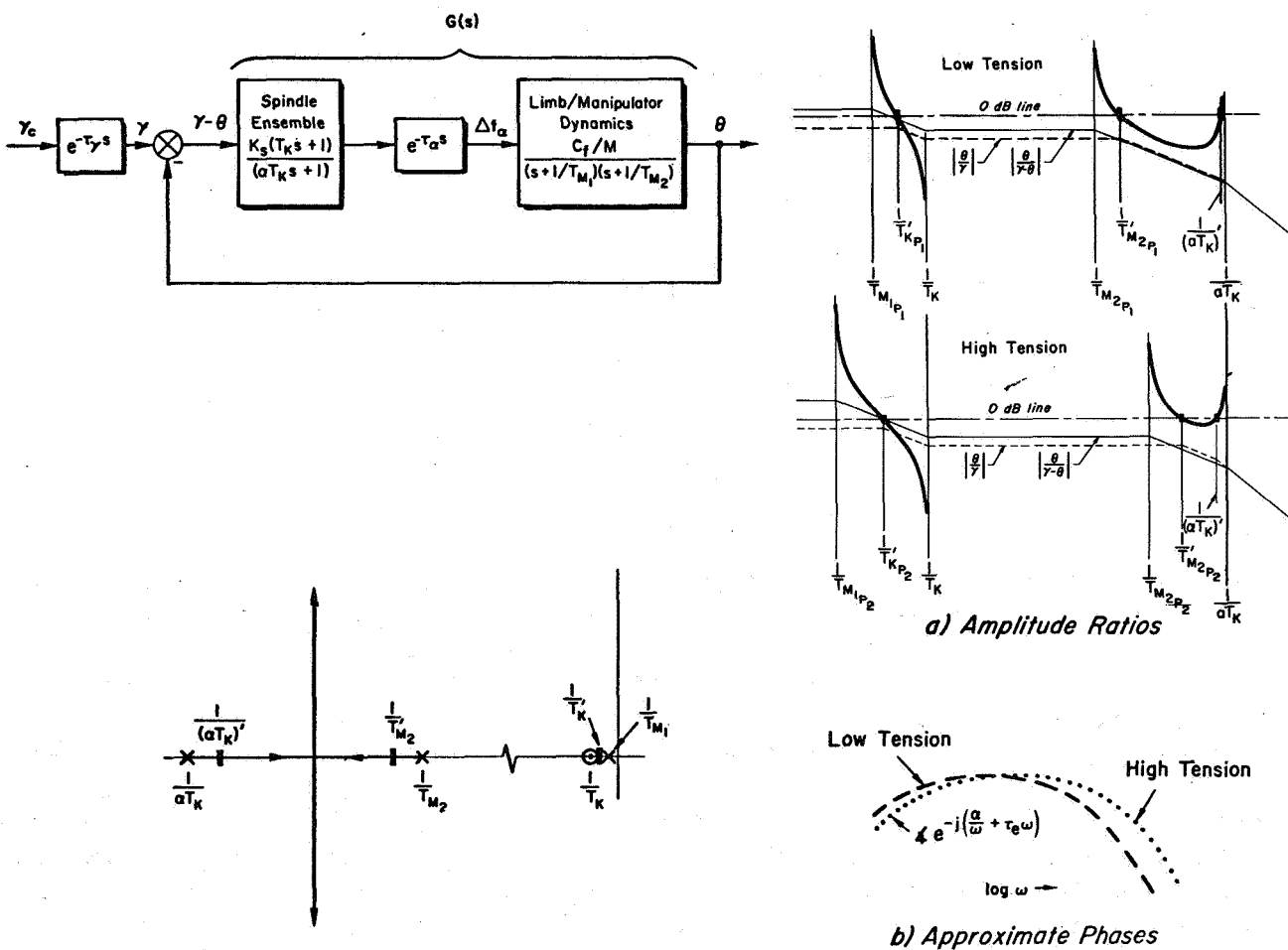


Figure 17.—Root loci of neuromuscular subsystem dynamics with two levels of tension.

(ref. 24 and appendix D of ref. 25). In both diagrams the small pure time delay, τ_α , within the loop is neglected. These plots show that as gain is increased the low frequency muscle/manipulator root $1/T_{M_1}$ approaches the lead zero $1/T_K$ of the spindle, while the high-frequency muscle root $1/T_{M_2}$ and the spindle lag $1/aT_K$ approach one another, rendezvous, and break into a second-order pair. For a particular loop gain, for instance, that shown by the zero-dB line on the amplitude-ratio portion of figure 17, the closed-loop dynamics for a gamma command input have the form

$$\frac{\theta}{\gamma_c} \doteq \left(\frac{K_{dc}}{1+K_{dc}} \right) \frac{(T_K s+1) e^{-\tau_\gamma s}}{(T'_K s+1) (T'_{M_2} s+1) [(aT_K)' s+1]} \quad (18)$$

Here K_{dc} is the dc gain of the open-loop system, τ_γ is the net delay in the gamma motoneuron pathways from the presumed point of command insertion in higher centers to the spindles, and the primes over the denominator time constants indicate that the constants are closed-loop quantities. Asymptotic plots of the open-loop $|\theta/(\gamma-\theta)|$ and closed-loop $|\theta/\gamma|$ are also given on the amplitude-ratio plots of figure 17. On these the open-loop poles and zeros coincide with breakpoints of the solid-line asymptotes, whereas the closed-loop poles and zeros are indicated by breakpoints on the dashed or dotted asymptotes.

Consider now the effect of tension variation on the neuromuscular system dynamics indicated in the plot. For a typical low-tension condition (see the top amplitude ratio and related phase plot), defined by the steady-state tension P_1 the muscle poles are located at $-1/T_{M_1 P_1}$ and $-1/T_{M_2 P_1}$. The closed-loop system will be given by

$$\left[\frac{\theta}{\gamma_c} \right]_1 \doteq \left(\frac{K_{dc}}{1+K_{dc}} \right) \frac{(T_K s+1) e^{-\tau_\gamma s}}{(T'_{K P_1} s+1) (T'_{M_2 P_1} s+1) [(aT_K)' P_1 s+1]} \quad (19)$$

or, in the midfrequency range,

$$\left[\frac{\theta}{\gamma_c} \right]_1 \doteq \left(\frac{K_{dc}}{1+K_{dc}} \right) \left(\frac{T_K}{T'_{K P_1}} \right) e^{-j(\alpha_1/\omega + \omega \tau_{e_1})} \quad (20)$$

This last expression corresponds to the phase description used in the approximate model described in the section entitled "Human Operator Data Indicating Neuromuscular System Effects."

Presume now that the steady-state tension is changed by modifying γ_0 ; then the open-loop plot changes to the amplitude ratio and phase labeled "high tension." On these, the low- and high-frequency muscle factors are decreased to $1/T_{M_1 P_2}$ and increased to $1/T_{M_2 P_2}$, respectively (for simplicity, the spindle lead $1/T_K$ and lag $1/aT_K$ are assumed to be unmodified by the tension change). With the midfrequency gain unchanged, the closed-loop dynamics are modified significantly. The low-frequency lag/lead is more widely spaced, giving rise to a larger low frequency phase lag (i.e., $\alpha_2 > \alpha_1$) as seen in midfrequencies; and the high frequency phase lag is substantially reduced (i.e., $\tau_{e_2} < \tau_{e_1}$). This demonstrates that the extremely

simple neuromuscular system model described in figure 17 is qualitatively compatible with the data summarized in the section entitled "Human Operator Data Indicating Neuromuscular System Effects" and the physiological "component" characteristics of the section entitled "Description of the Components." These results lend strong support to the neuromuscular system structure and connections presumed in figure 17. They also provide detailed functional roles within this system for the behavioral characteristics manifested by the muscle/manipulator and muscle spindle components when operated as separate entities,

While the behavior of the model closely approximates the original experimental data, it is essential to note some caveats. These are of two kinds—quantitative difficulties and oversimplified descriptions. For the first, suffice it to say that no data currently exist on spindle and muscle/manipulator ensembles for the motions for which the basic α and τ_e data are derived. Consequently, data on component ensembles cannot yet be shown to yield the closed-loop results already observed. Data of this nature for the muscle groups should be fairly simple to obtain, although those for the spindle ensembles are not presently within the possibility of direct experimental validation. If, however, the neuromuscular system structure presented here is assumed a priori, then in situ spindle implicit "measurements" can be obtained by suitable operations with the overall system and muscle subsystem data. The second point, that of oversimplification, is by no means as difficult to alleviate. For example, the muscle contains a series elastic component which makes the muscle/manipulator combination a third-order system. For spring-restrained low-inertia manipulators this component is much larger than the other elements and thus has been neglected in this discussion. However, for other restraints it can be important; this is currently being investigated. In addition, we are working on alternative hookups of the neuromuscular system components.

REFERENCES

1. Elkind, J. I.; and Forgie, C. D.: Characteristics of the Human Operator in Simple Manual Control Systems. IRE Trans. on Automatic Control, vol. AC-4, May 1959, pp. 45-55.
2. McRuer, Duane; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems: Theory, Models, and Experiments with Controlled Element and Forcing Function Variations. AFFDL TR-65-15, July 1965.
3. Young, Laurence R.; and Stark, Lawrence: Biological Control Systems—A Critical Review and Evaluation, Developments in Manual Control. NASA CR-190, 1965.
4. Jex, H. R.; McDonnell, J. D.; and Phatak, A. V.: A "Critical" Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time. Part I: Theory and Experiments with a First-Order Divergent Controlled Element. NASA CR-616, 1966.
5. Okabe, Y.; Rhodes, H. E.; Stark, L.; and Willis, P. A.: Transient Responses of Human Motor Coordination System. MIT Res. Lab. Elec. quar. progress rept. no. 66, 1962, pp. 389-395.
6. Houk, James Charles, Jr.: A Mathematical Model of the Stretch Reflex in Human Muscle Systems. Master of Science Thesis, MIT, 1963.
7. McRuer, Duane: Remarks on Some Neuromuscular Subsystem Dynamics. IEEE Trans. on Human Factors in Electronics, vol. HFE-7, no. 3, Sept. 1966.
8. Crowe, A.; and Matthews, P. B. C.: The Effects of Stimulation of Static and Dynamic Fusimotor Fibres on the Response to Stretching of the Primary Endings of Muscle Spindles. J. Physiol., vol. 174, 1964, pp. 109-131.
9. Crowe, A.; and Matthews, P. B. C.: Further Studies of Static and Dynamic Fusimotor Fibres. J. Physiol., vol. 174, 1964, pp. 132-151.

10. Stark, L.: Neurological Organization of the Control System for Movement. MIT Res. Lab. Elec. quar. progress rept. no. 61, Apr. 15, 1961, pp. 234-238.
11. Wilkie, D. R.: The Relation Between Force and Velocity in Human Muscle. J. Physiol., vol. 110, 1950, pp. 249-280.
12. Hill, A. V.: The Heat of Shortening and the Dynamic Constants of Muscle. Proc. Roy. Soc., 1938, B 126, 136.
13. Bigland, Brenda; and Lippold, O. C. J.: The Relation Between Force, Velocity, and Integrated Electrical Activity in Human Muscles. J. Physiol., vol. 123, 1954, pp. 214-224.
14. Wilkie, D. R.: The Mechanical Properties of Muscle. Brit. Med. Bull., vol. 12, no. 3, 1956, pp. 177-182.
15. Granit, Ragnar: Neuromuscular Interaction in Postural Tone of the Cat's Isometric Soleus Muscle. J. Physiol., vol. 143, 1958, pp. 387-402.
16. Swett, John E.; and Eldred, Earl: Distribution and Numbers of Stretch Receptors in Medial Gastrocnemius and Soleus Muscles of the Cat. The Anatomical Record, vol. 137, no. 4, Aug. 1960, pp. 219-227.
17. Matthews, P. B. C.: Muscle Spindles and Their Motor Control. Phys. Rev., vol. 44, 1964, pp. 219-288.
18. Barker, David: The Structure and Distribution of Muscle Receptors. Presented at the Symposium on Muscle Receptors, Golden Jubilee Congress, Univ. Hong Kong, Sept. 11-16, 1961.
19. Granit, Ragnar, ed.: Muscular Afferents and Motor Control. Proc. of First Nobel Symposium, June 1965, John Wiley and Sons, New York, 1966.
20. Lippold, O. C. J.; Nicholls, J. G.; and Redfearn, J. W. T.: Electrical and Mechanical Factors in the Adaptation of a Mammalian Muscle Spindle. J. Physiol., vol. 153, 1960, pp. 209-217.
21. Lundberg, A.; and Winsbury, G.: Selective Adequate Activation of Large Afferents from Muscle Spindles and Golgi Tendon Organs. Acta Physiol. Scand., vol. 49, 1960, pp. 155-164.
22. Harvey, R. J.; and Matthews, P. B. C.: The Response of De-Efferented Muscle Spindle Endings in the Cat's Soleus to Slow Extension of the Muscle. J. Physiol., vol. 157, 1961, pp. 370-392.
23. Whitteridge, D.: The Effect of Stimulation of Intrafusar Muscle Fibers on Sensitivity to Stretch of Extraocular Muscle Spindles. Quarterly J. Exp. Physiol., vol. 44, 1959, pp. 385-393.
24. McRuer, Duane T.: Unified Analysis of Linear Feedback Systems. ASD-TR-61-118, July 1961.
25. Stapleford, Robert L.; Johnston, Donald E.; Teper, Gary L.; and Weir, David H.: Development of Satisfactory Lateral-Directional Handling Qualities in the Landing Approach. NASA CR-239, 1965.

23. Pupil Dilation as a Measure of Workload

R. O. Anderson and P. E. Pietrzak
Air Force Flight Dynamics Laboratory

N68-15924

In April 1965 Eckhard Hess (ref. 1) of the University of Chicago presented some very interesting pupilometric (pupil diameter) measurements of subject interest, emotion, attitude, and thought processes. Hess referred to the eye as an extension of the brain that is in plain sight for the psychologist or engineer to observe. It was speculated that the implementation of this measurement for pilot workload is similar and perhaps superior to such measures as electroencephalogram (EEG) and galvanic skin response (GSR).

Supported by this somewhat questionable speculation, an experiment was performed to determine the feasibility of pupilometrics as a measure of workload. With expediency the prime factor, equipment was assembled, experiments performed, data reduced and analyzed, and results written up in approximately 4 weeks.

The basic equipment is seen in figure 1. From right to left there is an 8-channel strip-chart recorder, an overgrown desk-top analog computer, cathode ray tubes (CRT) for task display, a side-stick controller, a 16-mm electric drive camera, and a subject. The second scope was set up for a side task to which the subject responded with a thumb switch in the left hand.

The experiment design consisted of implementing the STI "critical task" on the analog computer. The compensatory task error display was a horizontally moving dot on the CRT which the subject controlled with lateral movements of the control stick. The subject, a military pilot in the Canadian Air Force, was told to keep the dot centered on the scope with a minimum of error. There was no disturbance input, and therefore the system output was a result of pilot actions only.

Four specific experiments were performed:

- (1) Time varying λ tracking
- (2) Fixed λ tracking (various λ values)
- (3) Fixed λ tracking with side task
- (4) Error observation only, no other tasks

In these experiments λ is the value of the controlled element pole in the right half plane, and increasing λ represents an increasingly more "difficult" task.

The data, consisting of film strips recorded at about 10 frames per second, were reduced by displaying the strips on a microfilm reader and measuring the pupil diameter with an optical micrometer. The results of this preliminary investigation can be seen in figure 2.

The top trace is the error seen by the subject. The second record is the value of λ showing the two ramp rates. The second rate was initiated when the display error exceeded an error magnitude criterion. The bottom trace is R, the ratio of pupil diameter to iris diameter. This ratio was taken to account for head motion. In addition, ratio values for constant values of λ are superimposed at the time point corresponding to the ramp λ .

The main points that we can draw from this data run are:

- (1) A large variation in R (about 36 percent) over the 46-second run
- (2) A steady increase in R with increased λ and increased error activity
- (3) A rapid change in pupil dilation with large error excursions
- (4) A steady increase in constant λ values

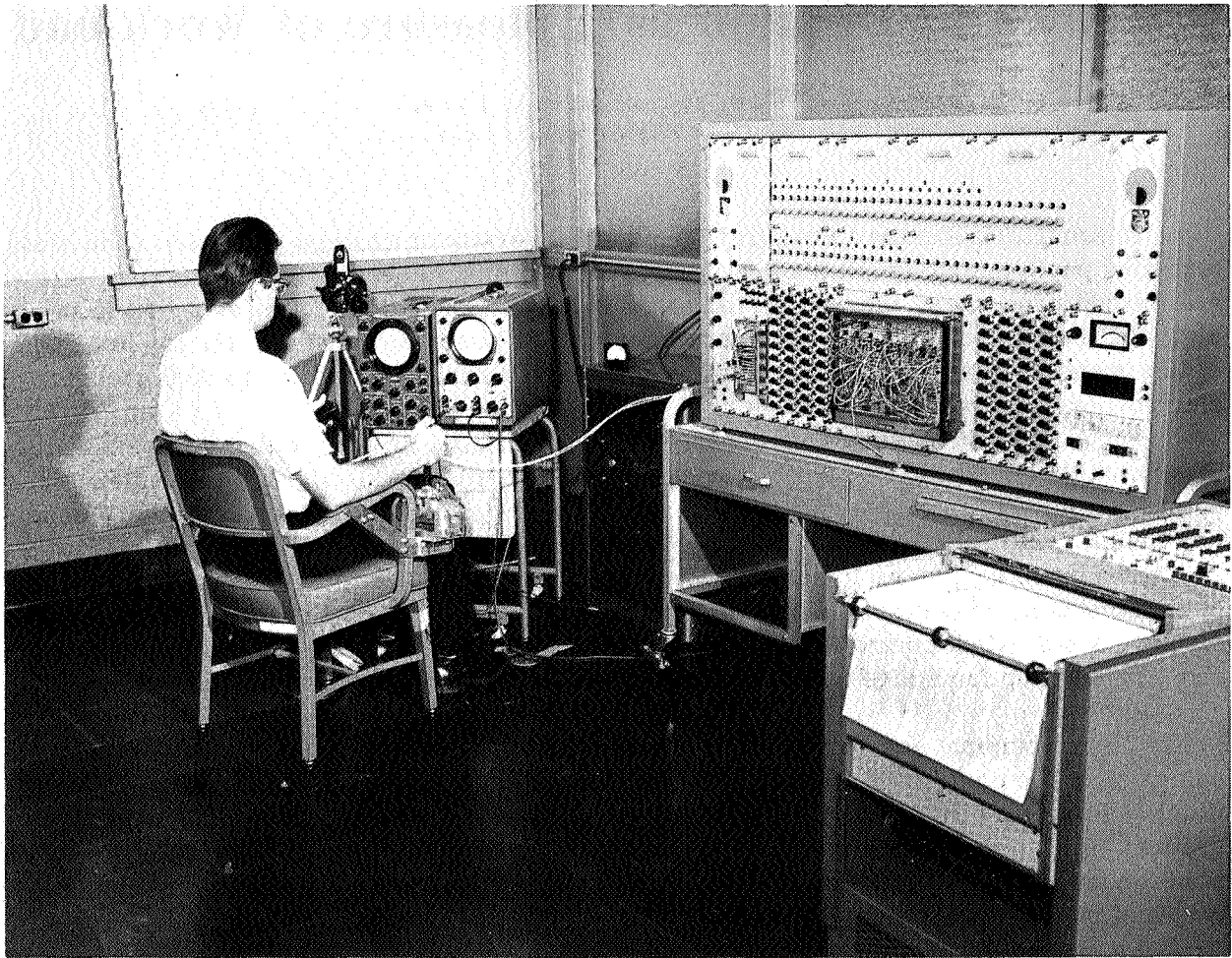


Figure 1.—Basic equipment.

The fluctuations in the **R** trace may be due to measurement errors or due to the "unrest" (noise) fluctuations reported by Stark.

In an attempt to determine if the pupil variations observed were the direct results of observing the display only and not unique to the combined observation-control task, a limited amount of data were collected where the subject viewed his error trace from a previous run and applied no control input. Figure 3 shows the last 20 seconds of a time varying **A** run and the same 20-second portion where the subject merely observed.

The film records indicated much more eye blinking and some eye tracking in the monitoring record. The tracking record showed a steady "stare." Although some correlation is apparent, a grainy film in the monitored run caused some data-reduction problems.

A short series of runs with a more conventional side task indicated a direct correlation of increased workload, measured in this manner, with pupil dilation. However the data population was very small.

Although the experiments reported here are certainly preliminary, the main objectives were met, and the following tentative conclusions can be drawn:

- (1) Pupil dilation is evident in certain manual tracking tasks of increasing "difficulty."
- (2) This dilation is correlated with the results of at least one, more or less conventional, workload measurement technique, and is also correlated with task "difficulty."

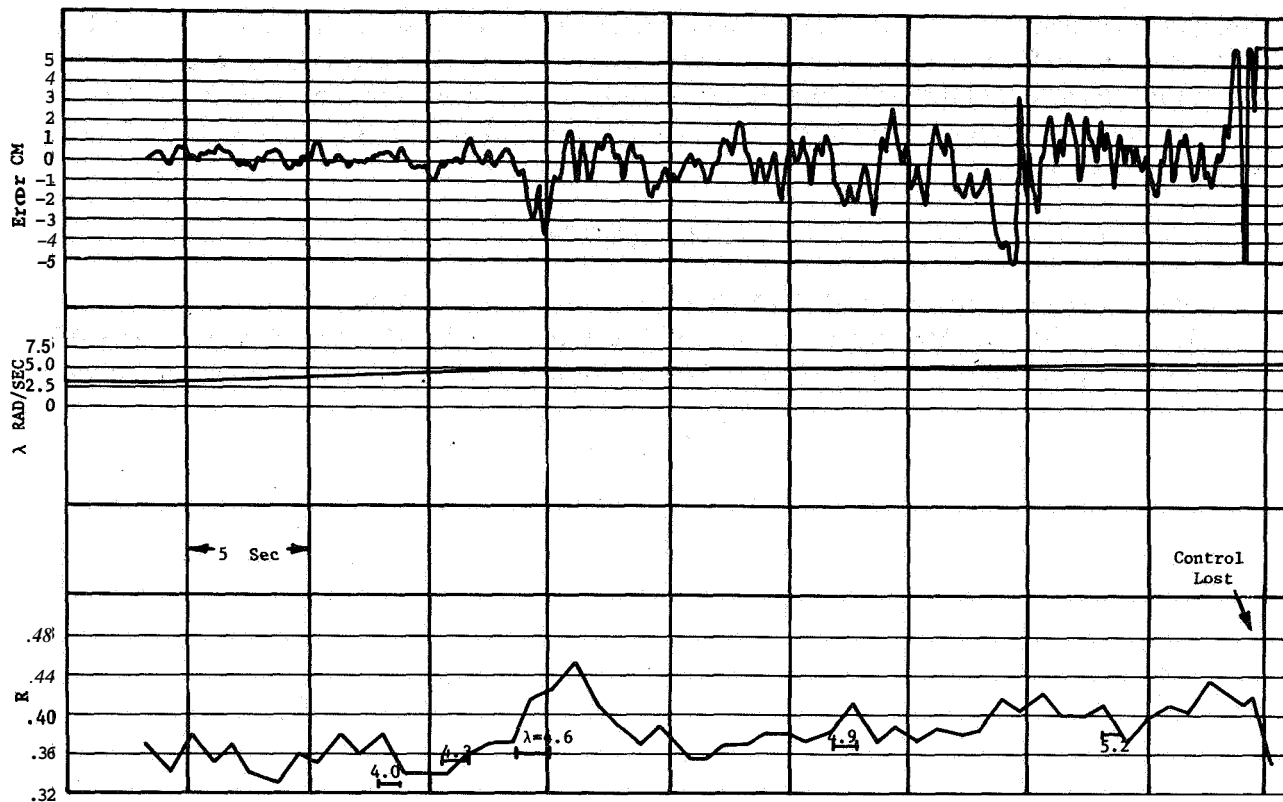
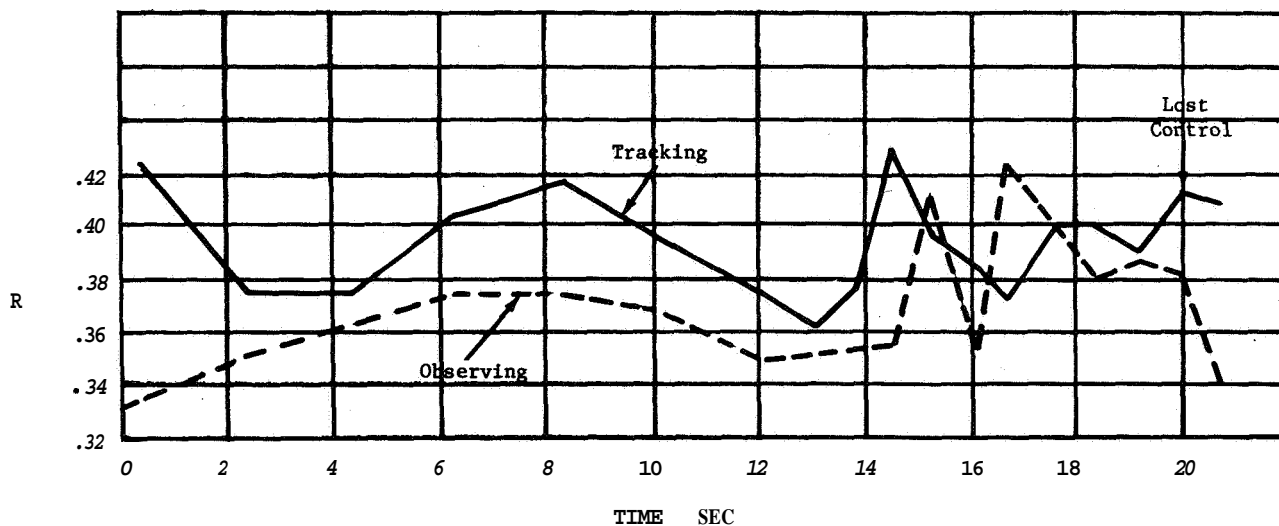
Figure 2.—Results of time-varying λ experiment.

Figure 3.—Pupil variations in tracking versus error observations.

(3) The exact cause of this effect is unknown, but it does not appear to be the sole result of error observation. That is, the phenomenon appears to be a result of stresses from the combined observation-control task.

Much remains to be done before pupil size variations can be used with confidence as a measurement technique for manual control workload, but the preliminary results reported

here are encouraging. Furthermore, the advantages of the method, if verified by further work, over side-task methods are clear with respect to rapid response, the degree of subject distraction from the main task, and the possibility of electronic on-line data reduction. A more detailed summary of the experiment and a general discussion of pilot workload can be found in reference 2.

REFERENCES

1. Hess, Eckhard: Attitude and Pupil Size. *Sci. Am.*, Apr. **1965**.
2. Westbrook, C. B.; Anderson, L. O.; and Pietrzak, P. E.: Handling Qualities and Pilot Work Load. **FDCC TM-65-5**, Air Force Flight Dynamics Laboratory, Sept. **1966**.

VIII. DECISION PROCESSES III

2000

2000

1

2000

2000

2000

2000

2000

2000

2000

2000

2000

24. "Inhibitory" Control: Concept for a First Model*

*John Lyman and Amos Freedy
University of California, Los Angeles*

N68 - 159251

Human-operator control of a multi-dimensional manipulative device is normally accomplished by a sequence of decisions which are based on the instantaneous state of the device. When devices such as remote manipulators and the human hand are controlled through multiple information channels, the state of the device is determined by the logical interconnection of the linked channels. The required information rate for minimal time control is a function of the initial and the desired final state of the device modulated by the interconnection among the information channels.

The optimum coordinated control of a given device depends on establishing the best control input-state relationship. This is normally accomplished by the operator.

If the statistics of the required movement pattern are available, an n-space probability function which is a function of the initial state may be generated. This probability function would contain the information for predicting the most probable next movement. A relationship can be established between the instantaneous logical interconnection of the information channels and the probability of a future state. When the interconnection can be preset continuously, the information rate required for changing a state would be inversely proportional to the probability of this state being the future state. Under such conditions the operator's decision load would be reduced, since he would only have to "inhibit" unwanted motions. A preliminary approach to the theory and implementation for such a control system is discussed.

In human-control loops for the operation of multidimensional manipulative devices, such as artificial arms and remote manipulators, the operator acts as an error generator. To direct a particular end result the operator generates an error signal for each dimensional movement, and vectorially adds their components continuously until the desired spatial state is achieved. Such an operative procedure places a decision load on the operator in proportion to the level of skill required for success. Attempts have been made in the past to free the operator from his decision load by mechanizing all or part of the decision process. Aiding and quickening are classic examples. Preprogramming a repertory of tasks, as in the computer-operated Case Institute arm-aid, has been another approach.

Although these approaches have proven useful, they have the major disadvantage of not being adaptive. In this paper, we present the concept of an adaptive system which makes use of the observed high autocorrelation of manipulator movement states during general task

*This research was supported in part by the Veterans Administration under Contract V 1005P-9779.

performance. In other words, there seems to be favored paths of movement in manipulator control, so that with the manipulator in a given state, certain future states are more likely than others. If the manipulatory control system could adapt to these favored paths and could drive the manipulator along them unless corrected, the decision load associated with the corrections might be less than that normally encountered by the operator.

Direct control by the operator plus control by an intervening adaptive device has led us to the concept which we have called "inhibitory control." In this kind of control system the driving error signal would be continuously generated by an adaptive processor as a function of the history of the movement pattern. The operator would be able to override the error signal given by the processor and correct any unwanted movement;

THE SYSTEM

A block diagram of a model of the "inhibitory" control system is shown in figure 1. If we assume that when the operator applies no input to the logic unit, the system $H(s)$ is driven by the output of the processor. The expected output E_n of the processor is given as:

$$E_n = f_k [p_{n-1}, \varphi_{vv}(t_1 t_2)_{n-1}]$$

where p is the state vector of the output device, φ is the autocorrelation of the velocity vectors, and the function f_k is variable over the boundary space of movement.

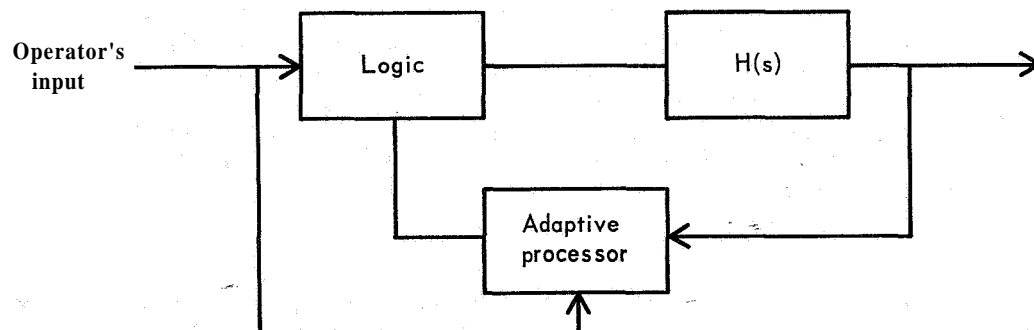


Figure 1.—Block diagram of a model of the inhibitory control system.

When the operator applies a correction input, the function f_k will readjust itself for a new value based on the input correction signal and the last value of f_k . For example, if E_n is given as a polynomial

$$E_n = K_k + L_k p + M_k \varphi + N_k p^2 + Q_k \varphi^2 + \dots$$

where $K_k=1, 2, 3, \dots$ marks the position in space.

Initially, the system is free running, at a state p and with velocity autocorrelation φ_{vv} . If the initial values of K_k , L_k , M_k , and N_k are not specified, the vectorial driving-input signal E_n will generate a random path. In order to direct the output device to a desired position the operator will have to add a correction signal D_n . On future trials for movements through the same state p , the processor will readjust the values of its constants to give a vector input consisting of a statistical average between its present value and past values. With repeated trials the values of the constants will converge to an optimum value in such a

way that E_n will approach an expected value desired by the operator. Correspondingly, the number of corrections required by the operator will be reduced with repeated trials. Thus the machine will reduce the operator's decision burden.

The inhibitory system can be realized as shown in figure 2. The forward loop consists of an input coder, the manipulator transfer function, and a gate control.

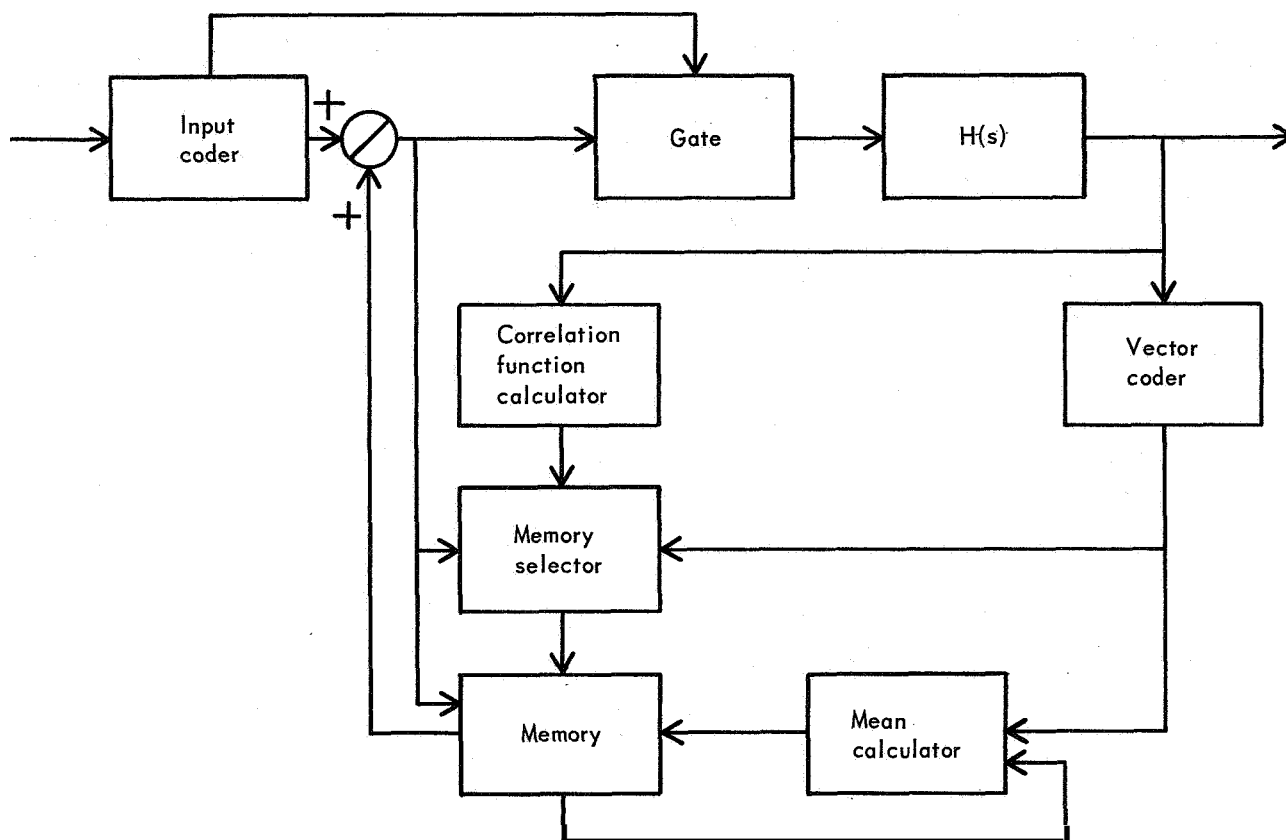


Figure 2.—The inhibitory control system.

The gate control sets the operation mode as "go" or "stop." The operator has binary control over each motion element of the system. The input coding unit integrates his binary input over time and produces negative or positive increments of input vector. When the system is set at the "go" mode by means of a particular signal of the multi-unit controller configuration, a control-input vector is given by the memory of the subsystem. The operator is able to change the memory control vector by introducing the vector D_n .

Each point in the bounded space of movement of the output device is mapped into each one of k -vector memory maps. Each spatial point of the memory map contains a mean value of all "experienced" past position vectors that were initiated on this point, when certain values of the velocity autocorrelation, position, and corrected operator input were accumulated. Thus, the selection of a particular memory map is identified by a configuration of the experienced past position vectors and operator corrections.

The autocorrelation of the velocity vector from the most recent path, the state of the device, and the operator's corrections are fed into the memory selector. The memory selector assigns a certain memory map for each condition. Given the necessary three conditions, the memory map which contains the required mean position vector for the future

state is selected with an increasingly biased probability. The probability that the proper memory map will be chosen can be generalized as:

$$P(E) = P(E/D_n) + P(E/\phi_{VV}) + P(E/p)$$

where p is the state vector.

Upon selection of the proper memory the output device will follow a pattern of movement which is the function of all previous paths when the same initiating conditions were given. Deviation from this path would require an additional input by the operator, which would add to the statistical average of the previously stored pattern of movement. A low correlation between the pattern of movement selected and the required path will result in a search for a memory map of a higher correlation.

It is assumed that the dynamics of the system will be predetermined for man-machine stability in terms of rise times, motion rates, and damping characteristics. We believe that this approach to manipulator control in which a once-initiated, ongoing system movement is a function both of probability bias based on experienced system function and voluntary operator correction is similar in principle to the learning of skilled movements by living systems. By providing a means to externalize partially the detailed decisions required for skilled movements, the operator can achieve more refined control of the output system. Through continued experience with his particular performance goals, the operator will, in effect, shape the plant function, combined with himself, into an integral unit.

BIBLIOGRAPHY

- Andrew, A. M.: Learning Machines. Symposium on the Mechanization of Thought Processes (Teddington, England), H. M. Stationary Office, Nov. 1958.
- Bekey, George: An Investigation of Sampled Data Models of the Human Operator in Control System. Tech. Rep. No. ASD TDR 62-36, Feb. 1962.
- Musses, C. A.: Aspects of the Theory of Artificial Intelligence. Plenum Press, 1962, pp. 1-7.
- Nachshon, A.: Skill Acquisition in Three Dimension End-Point Control. M. S. Thesis, Univ. Calif., Los Angeles, June 1965.
- Ullmann, J. R.: Cybernetics Models Which Learn Sensory Motor Connection. Med. Elec. and Bio. Eng., Jan.-Mar. 1963.
- Wijnschenk, M. J.: Engineering Evaluation and Preliminary Studies of the Case Research Arm Aid. Rep. no. EDC 4-64-3, Case Inst. Tech., 1964.

25. Supervisory Control of Manipulation^S

T. B. Sheridan and W. R. Ferrell
Massachusetts Institute of Technology

N68-15926

Several types of remote master-slave manipulators have been built for the Atomic Energy Commission (e.g., by Goertz (ref. 1) and by Mosher (ref. 2)); and several computer programs have been demonstrated through which mechanical hands, equipped with simple visual scanners and touch sensors, can carry out relatively specific manipulation tasks such as stacking blocks (e.g. Ernst (ref. 3) and M. L. Minsky of M.I.T. in a personal communication). While the arbitrary extension in space of man's combined visual, tactile, and motor capabilities promises much, our present understanding of human control capability in this context—not the hardware—poses severe constraints.

Earlier studies by the authors (refs. 4 and 5) indicated that the human can operate through a long transmission delay, provided he moves in a succession of open-loop moves and waits for feedback. Long delays necessitate long task-completion times if man remains in the loop. Consequently, it is evident that it would be better if man could remove himself from the immediate control loop and serve as a supervisor by intermittently setting subgoals for an otherwise autonomous remote system of servos, hands, and a small computer. This is illustrated in figure 1. The approach described in this paper treats the human operator as a strategic planner and is in contrast to the "inhibitory control" approach described by Lyman and Freedy in paper 24 of this volume. In that study the manipulator is conditioned by experience to move in certain patterns, and the human operator's task is to override manually and correct its errant moves.

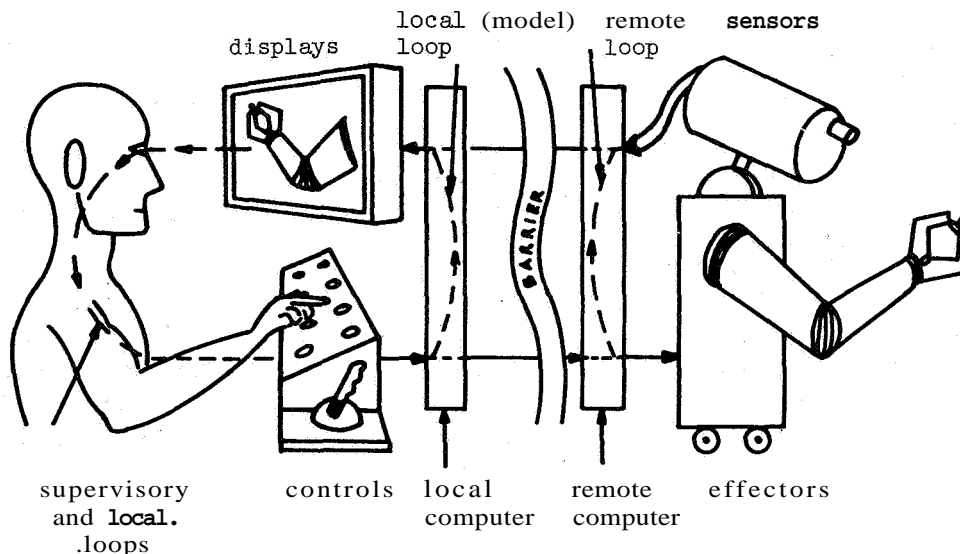


Figure 1.—Schematic diagram of supervisor-controlled remote manipulation system.

^SThis work was sponsored by the National Aeronautics and Space Administration under Grant NsG 107-61.

EXPERIMENTS

Recent investigation of human capability in the context of the present paper has included both empirical and theoretical approaches. These approaches have included the following type of experiments:

(1) Human controls computer-simulated three-degree-of-freedom manipulator and observes computer display

(2) Human controls actual seven-degree-of-freedom mechanical manipulator with typewriter controls

(3) Human controls actual mechanical manipulator with seven-degree-of-freedom joystick control plus several push buttons

(4) Human controls the modified carriage of an X-Y plotter through a joystick and K/s dynamics in order to push a nickel around on a plane.

Some analyses have also been made involving the application of dynamic programming and graph theory to the formal description and automation of aspects of the supervisory manipulation task.

McCandlish (ref. 6) performed an experiment in which the human operator adjusted knobs to make a pair of jaws move in two dimensions (scope display) in order to grasp an object, remove it from one hole, and place it in another hole (friction and gravity were included). In one case he performed the task with continuous rate control between knob position and jaw position. In a second case he could at several points command computer subroutines to take over certain segments of the task (see fig. 2). Figure 3 shows that under time delay this results in an improvement in number of commands used (an index of both time and signal transmission energy used) even for this simple task.

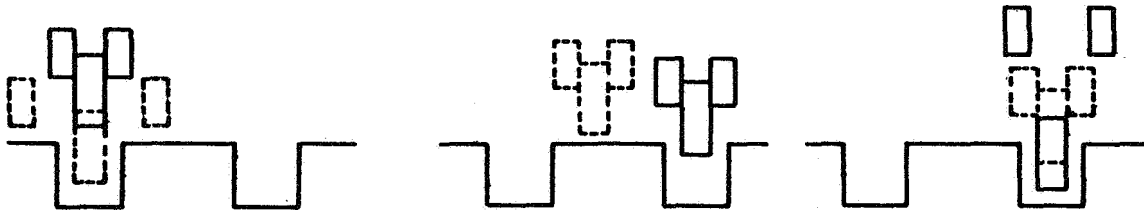


Figure 2.—Computer-generated display of simple manipulation task used in McCandlish experiment. In each of the three views of the jaws, manipulated object, and two-hole environment, the dotted lines represent the start and the solid lines the end of an available subroutine.

Rarich (ref. 7) and J. D. Barber of M. I. T. (unpublished results) have equipped a standard AMF-8 master-slave manipulator with stepping motors and interfaced this with a PDP-8 computer. In their system, the human operator types his commands as shown in figure 4. The manipulator hand is equipped with simple touch sensors (fig. 5) in order to provide feedback within the lower-level automatic-control loop of figure 1. Rarich and Barber have developed several kinds of programs by which the human operator may issue commands to the manipulator. Their command modes are listed in order of sophistication as follows:

(1) Rate control—YLZO (motor Y, direction L, speed 20) plus stop, jaw open, jaw closed.

(2) Increment control—YL20 (motor Y, direction L, distance 20) plus jaw open, jaw closed.

(3) Configuration name—User can name position configuration and at later time have manipulator resume that position merely by calling that name. Can also erase names.

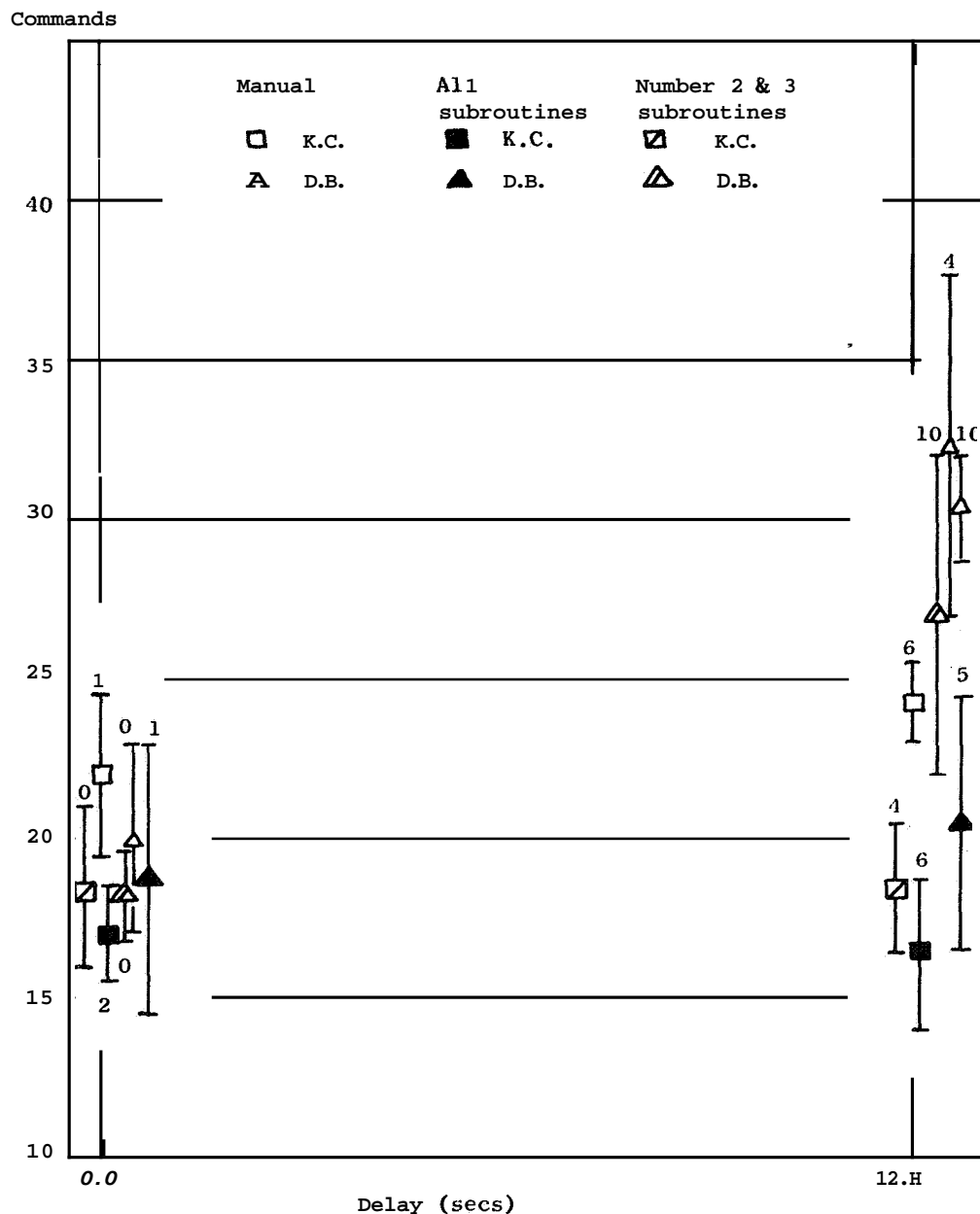


Figure 3.—Effect of automatic subroutine use on commands used. The figures at the end of a standard deviation bar indicate the number of errors made per 10 completed runs.

(4) Touch sensor satisfy—User can specify one of a combination of touch sensors and indicate on or off. Computer will decide direction to move and will stop only when that combination of touch sensors is turned on or off.

(5) Emergency instructions, calibration instructions, etc.

An advantage of a teletypewriter controller is that it easily permits the specification of conditional or branching commands such as

DISPLACE HAND LEFT UNTIL 40 OR TOUCH 2 OR 3



Figure 4.—Human subject controlling mechanical hand to pick up coffee cup through use of symbolic commands.

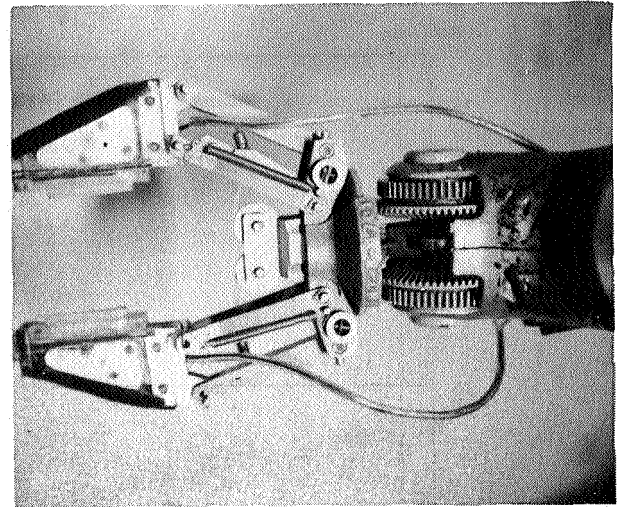


Figure 5.—Detailed view of manipulator hand showing wrist articulation and touch sensors on distal and palmer surfaces of two jaws.

IF 40 STOP, OTHERWISE ROTATE HAND RIGHT UNTIL 90 OR TOUCH ANY.
IF 90 STOP, OTHERWISE 527.

Note that while in a sense the above is a computer program, it is also a direct command from a human to a mechanical device to move in three dimensional space. Some other types of branching commands are illustrated in figure 6.

Some experiments have been conducted by Verplank (ref. 8) using a shoulder-mounted controller (fig. 7); the degrees of freedom of the controller are geometrically isomorphic with

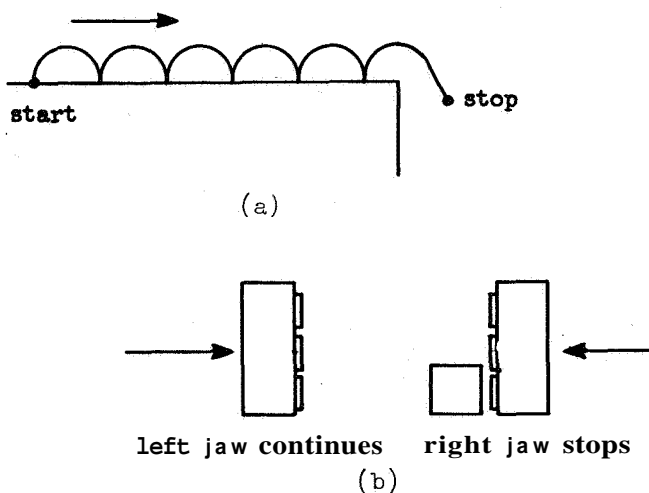


Figure 6.—Command subroutines. (a) Test surface in given direction for corner; (b) grasp object without displacing it.

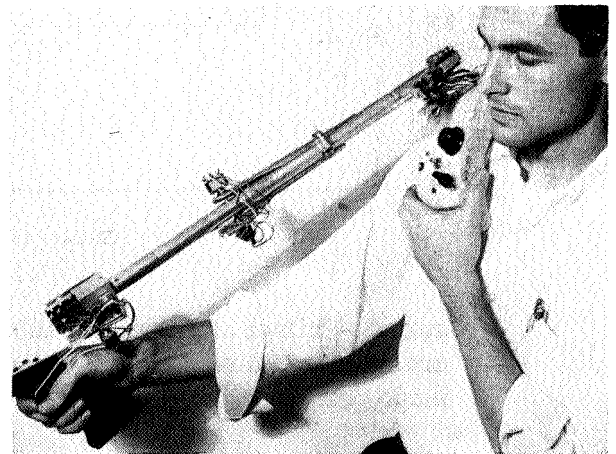


Figure 7.—Analogic controller devices. In each of seven degrees of freedom (which correspond to manipulator) operator can operate a three-positional (plus, zero, minus) switch. Knobs on the shoulder piece are used for switching computer modes (increment, rate, etc.).

the manipulator, but the controller offers only on-off-reverse control in each degree of freedom. For the increment and rate command modes, this device permits more speed and greater dexterity than the typewriter controller.

TOWARD A THEORY

We have felt obliged to explore the differences and similarities between conventional tracking and vehicle control and what we call remote manipulation or tool-object control. One rather important distinction appears in the latter when the controlled tool at times becomes disengaged entirely from the manipulated object, the position of which is the final output of the system.

Figure 8 illustrates several varieties of manipulation. At the top, for contrast, is vehicle control. This and the other three control systems we assume to be "self-paced." Self-pacing means that the "input" is a desired position or time-ordered set of positions which the object must assume, but the desired positions do not necessarily have unique times associated with them. Direct manipulation (b) is different in that the handle (tool) is now

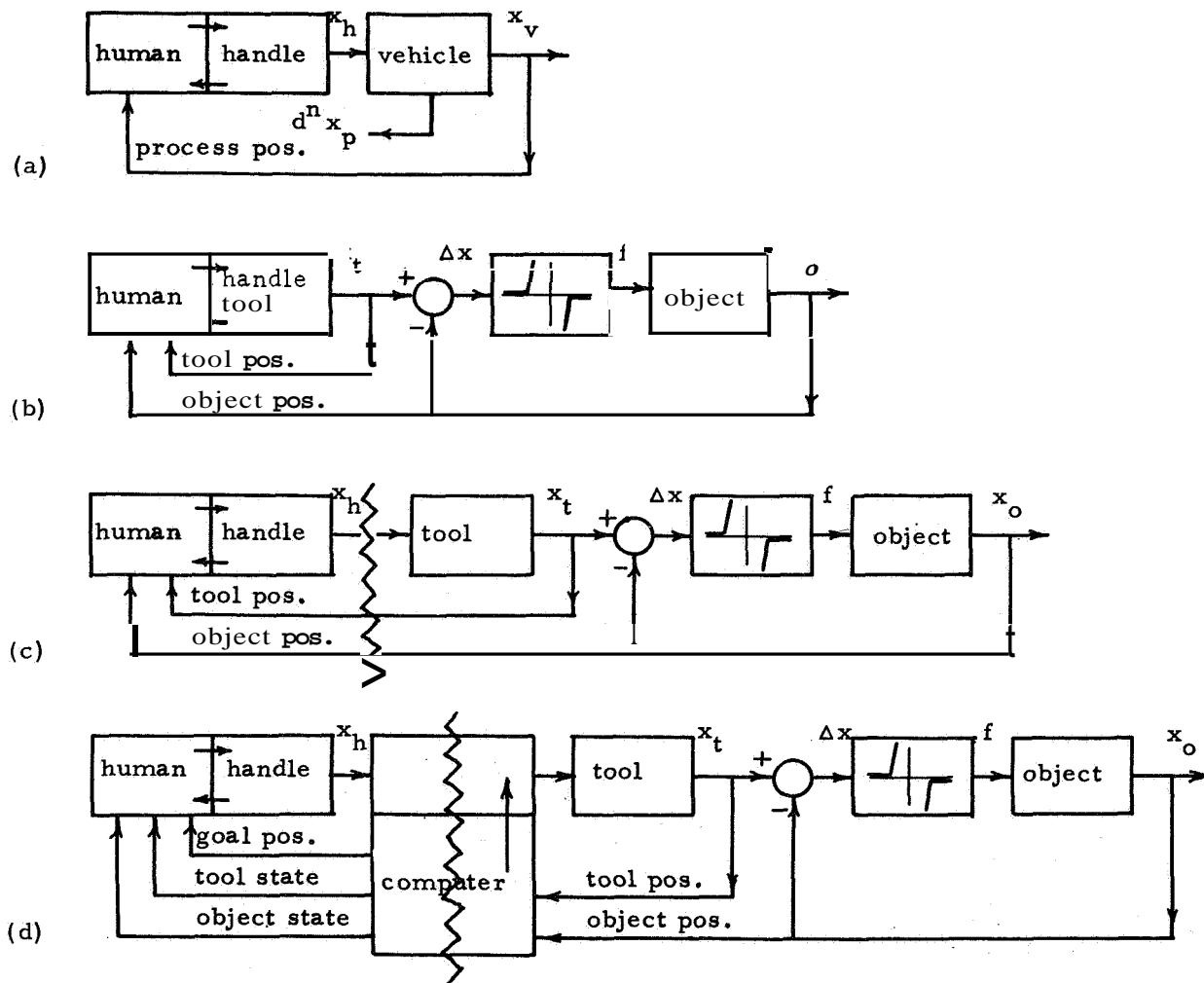


Figure 8.—Comparison of vehicle control and several varieties of manipulator systems. (a) Vehicle control; (b) direct manipulation; (c) remote manipulation; (d) supervisory manipulation.

observed by the human directly but in relation to the object position. When the tool approaches a fixed distance from the (center of the) object, force is applied to the object. This very important nonlinearity is fundamental to all tool-object manipulation and is represented schematically in (b), (c), and (d). Remote manipulation (c) is similar to direct manipulation except that some kind of a barrier (zig-zag line) is imposed between the handle and the tool; the barrier may take the form of time, space, noise, etc. in either forward or feedback loops. For convenience we now regard the tool as a separate dynamic element from the human-handle combination. Note the presence of what, in vehicle control parlance, is an "inner loop": the tool must be controlled in order that subsequent control of the object may be effected. Finally, at (d) the human is uncoupled from the tool-object continuous-control loop. The human gives commands to a computer on his side of the barrier which encodes and transmits the commands to a computer on the remote side of the barrier, where the continuous control loop is closed and lower level decisions are made. For any of the three manipulation systems when the object is rigidly grasped by the tool or remote manipulator, tool and object become one body and the contact nonlinearity vanishes.

To investigate the nature of the nonlinearity which separates tool and object but which has no counterpart in conventional tracking (tugboat-barge operations, rendezvous and docking are examples of the former, not the latter), a simple experiment was set up with an X-Y plotter (fig. 9). The subject controlled the carriage using a joystick and simple K/s dynamics in both X and Y degrees of freedom. The task was repetitively to push and position a nickel from within one circle (slightly larger than the nickel) across the XY table to a position within a second circle, and then back. The XY carriage was equipped with three different pusher

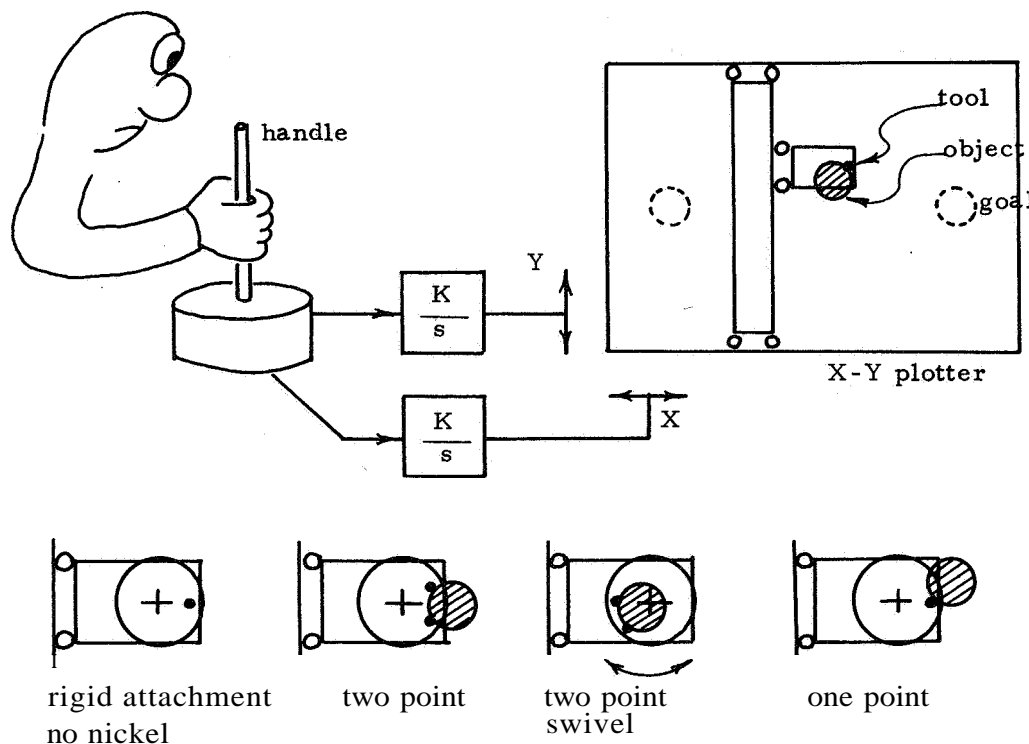


Figure 9.—Simple experimental demonstration of manipulation performance with different types of tool-object contact. Task was to position object (nickel) between dotted circles repetitively using various tool configurations.

mechanisms representative of three levels of security of grasp in manipulator hands (or prosthetic devices, etc.) :

- (1) Single-point contact
- (2) Two-point contact with a swivel, that is, limited ability to support moments
- (3) Two-point contact with rotational constraints

It is our tentative hypothesis that grasping an object usually involves a progression from (1) to (3). As an experimental control a fourth condition was to position a pointer with no nickel.

Results for four subjects for each of the four conditions above at each of two integration rates K are shown in figure 10. Points represent total completion times for five complete

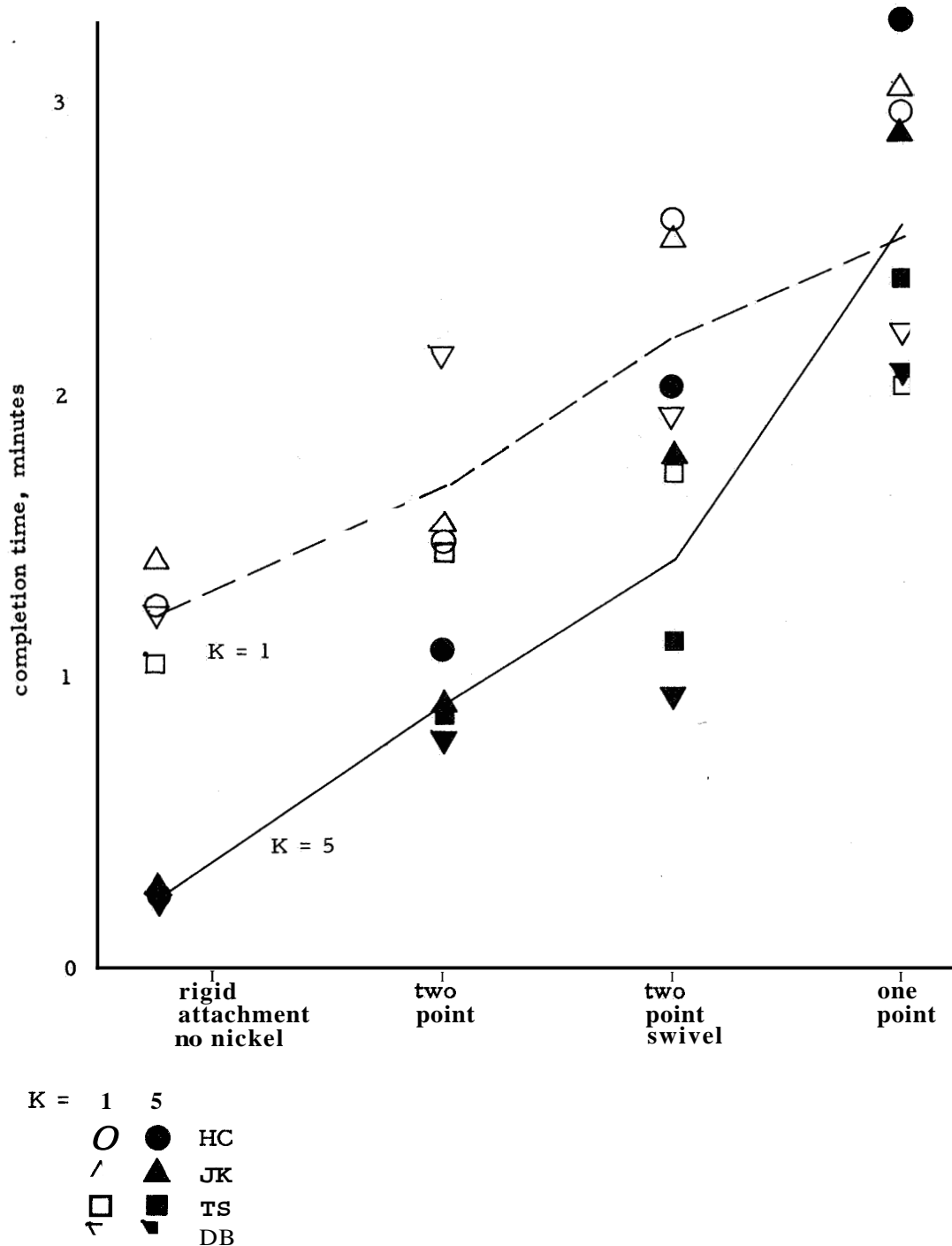


Figure 10.—Comparative results of nickel-pushing experiment.

cycles between points, and lines are corresponding averages across subjects. The consistent ordering of difficulty of the three types of pushers or contact elements (plus control) is obvious, as is the fact that high gain makes a big difference for simple tasks but is of no use when the more complex strategies are required. In fact, if the times for the control are subtracted out, leaving essentially the times for manipulation in and around the circles, it is seen that the high gain has a distinctly deleterious effect. Results suggest that the subjects would have done better had they simply eased up on the joystick in and around the target circles (which they could have done without adding undue noise). Training in particular strategies would probably improve performance, since intersubject variability in time is large for the one-point task where a variety of strategies was observed as compared with the two-point (nonswivel) task in which all subjects used the same strategy.

While a direct attack on the general problem of manipulation using optimal control or even a quasi-linear post-hoc description is unthinkable at this stage, some promising beginnings have been made to describe formally what occurs. For example, D. E. Whitney of

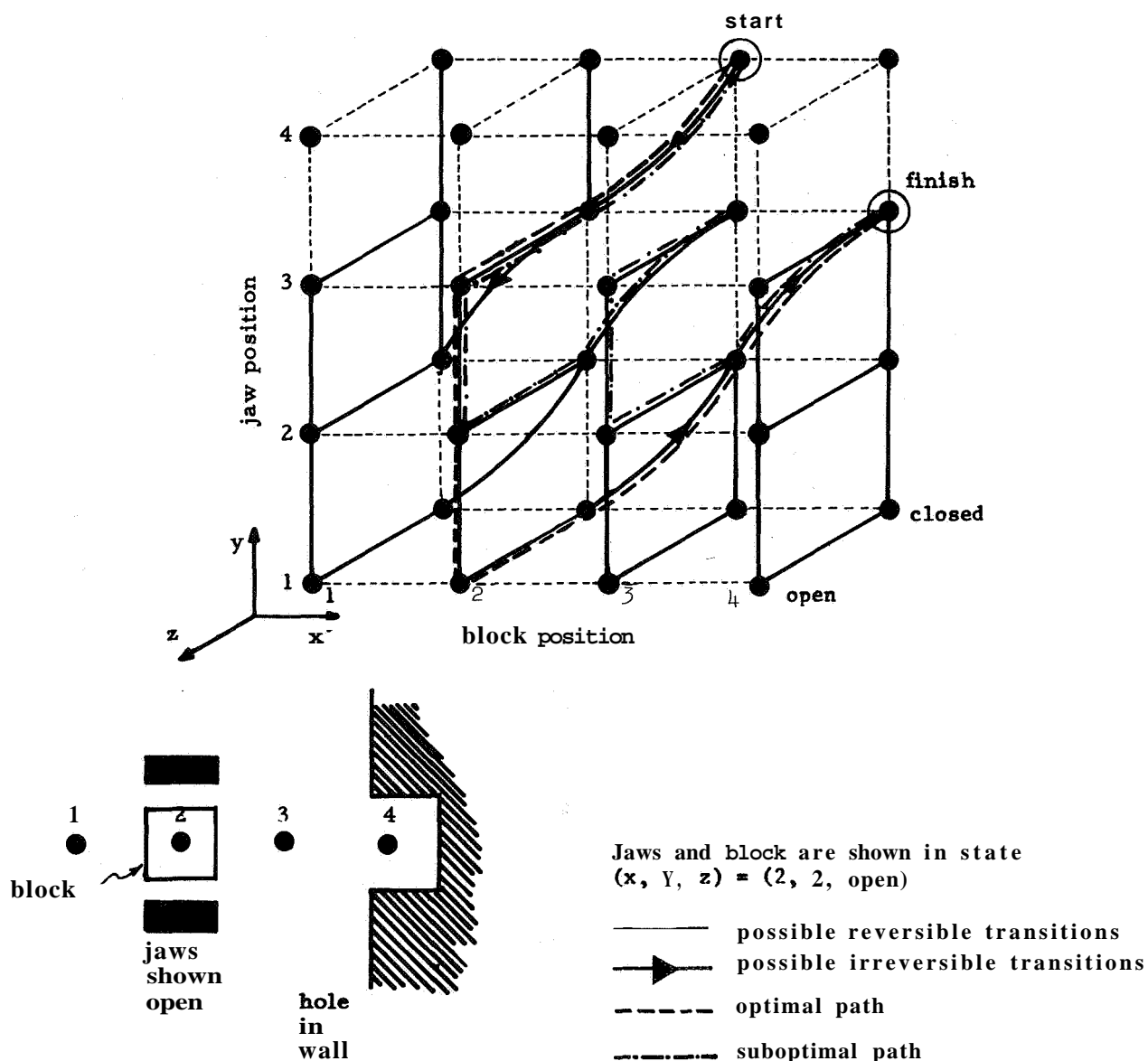


Figure 11.—State-space representation of simple task of placing block in hole.

M. I. T. has investigated a "manipulation state space," in which a point represents a particular combination of jaws and object positions. Because of the tool-object nonlinearity referred to above, transition between some adjacent states (edges between nodes of the state space graph, see fig. 11) are either impossible or irreversible. For example, the reader may observe from figure 11 that moving open jaws to 4 is impossible for any block position, indicated by the lack of any edges to $y=4$. Closed-jaw motion from $x=2$ and $y=1$ to poke the block into the hole at $x=4$ and $y=3$ is an irreversible diagonal motion. On the other hand motion from (1, 1, closed) toward (3, 3, closed) can only traverse the diagonal $x=y$, but nevertheless it is reversible. Considering the simple task of removing the jaws from a hole and placing an object in that same hole, one trajectory through state space (dashed line in fig. 11) can be shown to be better (shorter or otherwise less costly) than another trajectory (dash-dot line) for a given performance criterion (number of edges traversed in this case). Whitney has shown that, given a state space graph and knowing which edges can be traversed in which direction, it is possible, through a relaxation algorithm similar to dynamic programming (but possessing no monotonically advancing "stage" dimension), to find an optimal trajectory in a number of trials no greater than the number of states but usually much smaller. He has exercised this algorithm on the PDP-8 in several simple manipulation tasks.

REFERENCES

1. Goertz, R. C.: Manipulators Used for Handling Radioactive Materials. Human Factors in Technology, E. Bennett et al., eds., McGraw-Hill Book Co., 1963.
2. Mosher, R. S.: Industrial Manipulators. Sci. Amer., Oct. 1964.
3. Ernst, H. A.: MH1—A Computer Operated Mechanical Hand. Sc.D. Thesis, Dept. Elec. Eng., M.I.T., June 1961.
4. Sheridan, T. B.; and Ferrell, W. R.: Remote Manipulative Control with Transmission Delay. IEEE Trans. Human Factors in Electronics, vol. HFE-4, no. 1, Sept. 1963.
5. Ferrell, W. R.: Remote Manipulation with Transmission Delay. IEEE Trans. Human Factors in Electronics, vol. HFE-6, no. 1, Sept. 1965.
6. McCandlish, S. G.: A Computer Simulation Experiment of Supervisory Control of Remote Manipulation. Rep. DSR 9960-2, M.I.T. Eng. Proj. Lab., June 1966.
7. Rarich, T. D.: Development of SCM-1 A System for Investigating Performance of a Man-Computer Supervisory Controlled Manipulator. Rep. DSR 9991-3, M. I. T. Engr. Proj. Lab., May 1960.
8. Verplank, W. L.: Symbolic and Analogic Command Hardware for Computer-Aided Manipulation. S.M. Thesis, Dept. Mech. Eng., M.I.T., Jan. 1967.

PRECEDING PAGE BLANK NOT FILMED.

IX. ADVANCED MODELING TECHNIQUES

1

2

3

4

5

6

7

8

9

26. Stochastic Modeling of Human Learning Behavior

*Albert E. Preyss and Jacob L. Meiry
Massachusetts Institute of Technology*

768 - 15927

A stochastic model of human-learning behavior in a manual control task is described. Regulation of the state of a double integral plant to minimize the integrated absolute error is the operator's task. Subjects given this task were instructed to drive the process from an initial state to the null state using a two-position relay controller and a visual display.

A subject is conceptualized in the model as a sequential data-processing system. A sensor, a decision maker, and an effector are the three serially connected components making up the system. Each element requires a finite time either to process or transmit information, and thus a delay is incurred between the reception of the visual stimulus and the execution of a motor response. In agreement with known experimental evidence, this delay or reaction time is treated as a sum of random variables. A time delay of random duration thus completely describes the sensor and effector dynamics.

Decisions are made on the basis of an a priori estimate of the probability that the control polarity should be switched, given the current state of the plant. Patterns in the subsequent phase trajectory are used as evidence by the decision maker to revise a prior estimate. Bayes' theorem is the algorithm employed for the determination of the posteriori probability.

Behavior of this model is compared with subject behavior in the motor-skill experiment. The extent of the model's characterization of the time-varying random nature of human learning is brought out by this comparison.

Also discussed are the applications of the concept of this model to other manual control tasks.

Human learning behavior in a manual control task is the topic of this paper. A theory and model of motor skills learning are presented. Motor skills learning is given a stochastic interpretation by the theory. According to this interpretation, motor skills learning is a statistical revision-making process by which the human operator identifies a policy for the manual control of a dynamic process. This policy determines the limb movement he will make in response to a given sensory stimulus. Before this policy is identified, the human operator is uncertain as to which limb movement of a possible set of alternatives is the correct response to a given sensory stimulus. He must, nevertheless, respond to stimuli (no limb movement is also considered to be a response) while the manual control task is occurring. When he does, the theory postulates that his selection of a response alternative is based on his preferences for the alternatives in ranking them at the moment of choice, with these preferences being expressed as probabilities.

A control policy is identified, and therefore a manual control task is learned when the human operator resolves his uncertainty. That is, by this statistical revision-making process, he changes his preferences until all but one alternative response to a given sensory stimulus becomes improbable. Bayes' theorem is the proposed analog of man's algorithm for

revising his opinions, that is, for changing his preferences for alternatives. Using probabilities for the orderly expression of human opinion and representing statistical revision making by Bayes' theorem are concepts which characterize an application of Bayesian statistics for the probabilistic description of human-information processing. These ideas have been incorporated in the present theory to permit a complete mathematical treatment of a psychological phenomenon, the explanation of which is enhanced through quantification.

The theory postulates that the selection of response alternatives and the revision of preferences for response alternatives are functions of what is called the decision center of the human mind and that this decision center is one component of a single-channel information-processing system. Also included in this information-processing system are a sensor, which perceives the information upon which the decision center acts, and an effector, which executes the response decisions made by the center. An explanation of the operation of the sensor and the effector completes the description of human learning behavior provided by the theory.

A model of human learning behavior in a manual control task is readily constructed once the theory has been developed. The model is, in fact, a digital computer program which is obtained from a translation of the theory into machine language. There are a set of read-in parameters which govern the learning behavior of the program. These parameters correspond to human psychophysiological characteristics and can be adjusted to vary the individuality of the program. It is quite apparent that intersubject, intertrial response variability is present in any experiment involving the testing of humans. To ignore the factors which cause one individual's behavior to differ from another's or to differ from one time to another is to ignore what are frequently the major sources of variance in experimental data. Recourse to a stochastic model with randomly distributed parameters represents an attempt to account for these sources of variance.

THEORY AND MODEL OF LEARNING

Human operators in performing manual control tasks respond to sensory stimuli with limb movements. The development of our theory begins with the conceptualization of this psychophysiological phenomenon of response generation as a single-channel information-processing system. Following the presentation of this concept, we proceed to elaborate upon a description for each of the serially connected components comprising this information-processing system. There are three components to be considered: a sensor, a decision center, and an effector, and it is natural to treat them in this order, since it corresponds to the sequence in which we assume information is processed as it flows through the system. The operation of each component on the information transmitted to it is interpreted stochastically.

THEORY.—Before we begin the detailed development of the theory, we briefly outline our concept of how the human operator functions as a stochastic information-processing system. In our view, information, related to the state of the dynamic process being controlled and displayed to the human operator, is perceived by the sensor, quantized, and transmitted to the decision center. When the center is free to process new data, it accepts the most recently received sample of state information and decides upon a response to this stimulus. A decision is required because alternative responses to the same stimulus are possible. Stored in the memory of the decision center are the operator's preferences for the possible alternatives, and we express these as probabilities. Selection of a response is governed by a rule which takes the operator's preferences into consideration. Response decisions are then passed on to the effector for execution. Time elapses between the acceptance of a sample and the completion of the selection and between this moment and the execution of the response.

These intervals are treated as statistically independent random variables. During each of these cycles, the decision center may also take time out to revise the stored preferences before initiating the selection process, if it is deemed necessary. Revisions are based on the outcomes of previous response selections, a procedure which we refer to as the weighting of evidence and which we describe by an application of a set of ideas collectively called Bayesian statistics. Thus, the learning behavior of the system is characterized by a weighting procedure which revises preferences for possible response alternatives. Processing a revision adds to the delay between stimulus reception and response execution, and this increment is also treated as a random variable.

In this work we are concerned with the behavior of human operators who are learning how to regulate the state of a dynamic process by actuating a two-position relay controller. We assume that the process dynamics are second order and that the output of the process x is displayed to the operator. Further, the objective is to keep x nulled, and operator performance is scored on the basis of the integrated absolute value of x over the duration of a trial. In this **task**, the complete finger movement necessary to actuate the switch once is defined as an operator's response.

THE SENSOR.—In a manual control task, man's sensory apparatus provide him with the state information necessary to effect the closed loop control of a dynamic process. His perception of the displayed output, x , and of its rate of change, v , is subject to certain limitations. We know, for example, that angles can be resolved by the eye only to fractions of a degree and that the estimation of rates of change in stimulus dimensions is less certain than the estimation of the dimension itself. It is also known that sensory information is delayed in its transmission to the higher mental centers, so that the human operator's knowledge of the process' state is never current. Then, there is the question of how this information is coded for mental processing. Our description of the operation of man's sensory apparatus is an attempt to consider all of these factors in the simplest manner. Figure 1 shows a finite grid overlaying the state space of the dynamic process. It is assumed that sensory stimuli are categorized by the coordinates (m, i) of the mesh in which the process's state actually lies. We are saying, in effect, that the decision center, due to measurement errors and transmission delays, is certain of the current state of the dynamic process only to within the dimensions of a mesh and, furthermore, that even if the data could be resolved further, it would not be desirable to do so since processing times would be lengthened. We are, therefore, coding by quantizing.

THE DECISION CENTER

Subjective Probabilities.—When the decision center samples the state information transmitted by the sensor, it must use these data to decide upon a response. For the task in question this means choosing between the alternatives: to switch control polarity or not to switch. During the learning phase, the human operator does not know which of these two alternatives is correct. It may be said, therefore, that a state of uncertainty exists in his mind. Thus, before he can make a response, the human operator is forced to weigh each of the alternatives and, on the basis of some expression of preference, to select one. We propose the use of a probability as an expression of preference in the weighting of an alternative.

Hypothesized Control Policies.—A control policy for the regulation of a second-order dynamic process may be defined by specifying the switch curve. We will assume that stored in the memory of the decision center is a set of probabilities, $p \left[H_i(x_m) \right]$, for each of the $M \times N$ hypotheses,

$(H_i \ x_m) :$ The switch curve passes through the mesh, (x_m, v_i)

and that these probabilities are distributed so the conditions,

$$\sum_{i=1}^N p[H_i(x_m)] = 1 \quad (m=1, \dots, M) \quad (1)$$

are satisfied. A subset of these hypotheses such as the joint hypothesis,

$$[H_r(x_1), H_s(x_2), \dots, H_t(x_M)]$$

may be interpreted as the definition of a control policy, since it specifies the mesh, for every x_m , in which the control polarity should be switched.

Although it is possible to base switching decisions on these joint hypotheses and their probabilities, we propose a much simpler scheme. Before we do, though, recall the labelling in figure 1. The reason for drawing no distinction between the states $(x, -v)$ and $(-x, v)$ is that we are assuming that the switch curve is antisymmetric and, therefore, the same decision can apply to either state.

A Selection Rule.—The probabilities assigned to the hypotheses, $H_i(x_m)$, may be used in several ways to decide whether or not to switch control polarity when the sampled state is (x_m, v_j) . It is not likely that a single one of these methods will characterize the decision making of all human operators. Rather, one would expect the rule for the selection of an alternative to differ among individuals. Nevertheless, we will postulate a unique representation of the selection process in order to make the development of theory and model more tractable. The selection rule we propose depends on the probability,

$$Q_j(x_m) = \sum_{i=j}^N p[H_i(x_m)] \quad (2)$$

that the switch curve at x_m passes through a mesh whose velocity coordinate lies in the closed interval (v_j, v_N) . We refer to this as a switching probability and assume that

the selection of an alternative is a Bernoulli trial with probability $Q_j(x_m)$ of success, i.e.,

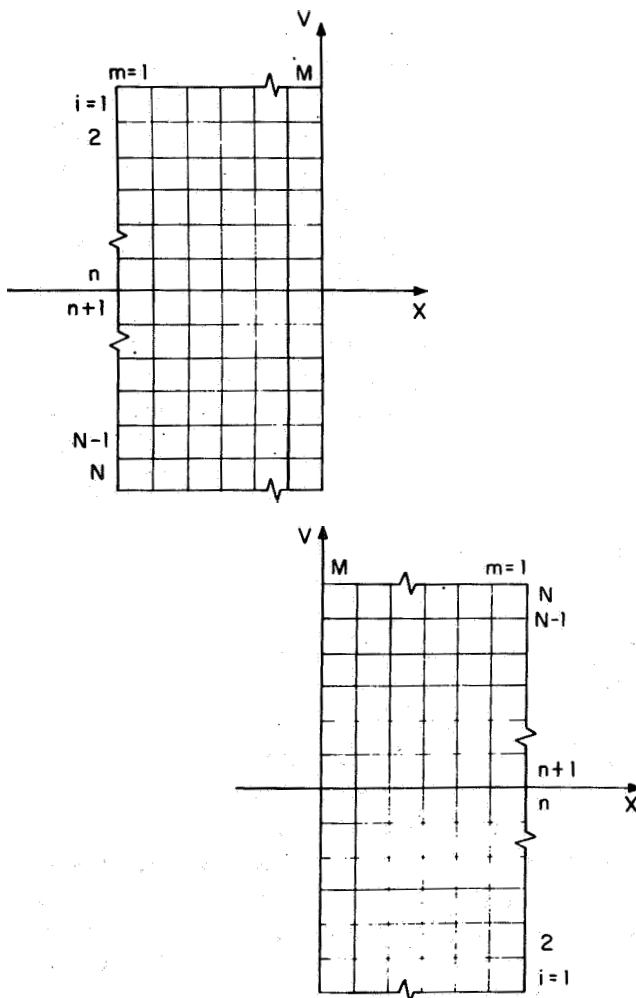


Figure 1.—Finite grid overlaying the state space of the dynamic process.

of switching. When we speak of switching with probability, $Q_j(x_m)$, we imply that the control polarity is opposite to the sign of x_m at the time of decision and that the switch will make the signs the same. In this case, $1 - Q_j(x_m)$ is the probability that the signs are kept opposite, i.e., no switch occurs. If at the time of decision the signs are already the same, the $1 - Q_j(x_m)$ is the probability that the center decides to switch control polarity to make it the opposite of sign (x_m) and in this case, $Q_j(x_m)$, is the probability of not switching.

Prior Probabilities.—Consider now what occurs on the first trial of a motor skill experiment. A subject has been briefed on the task he is to perform. On what does the human operator base his first response? If no clues have been provided by the briefing, any preference for a particular response must reflect a personal bias stemming from his past experience with similar or related tasks. Or, a subject might make a guess at what the dynamics of the process are and thereby be favorably disposed toward one control policy. Another subject may have very little experience with manual control tasks and may be initially inclined to treat the possible alternatives as equally likely candidates. Whatever his background, a subject's initial beliefs, those which he "brings with him," are expressed by the probabilities stored in the decision center's memory at the beginning of the experiment. These are called prior probabilities, and a decision to respond for the first time is based on them.

Revising the Prior Probabilities.—In order to learn a psychomotor task, the human operator must resolve his uncertainty as to the location in phase space of the switch curve. He may wish, therefore, to revise his opinions and express some other preference for the possible alternatives. A revision of opinion can be treated as a change in the prior probabilities $p[H_i(x_m)]$. In the terminology of statistics the revised opinion is commonly referred to as a posterior probability. Information used for the purpose of revising an opinion shall be called evidence E . Whatever the form of this evidence, the subject's use of it can be thought of as a weighting of the prior probability. Such a weighting may be represented symbolically in the following way:

$$p'(H_i) = w_i(E)p(H_i) \quad (3)$$

where the prime denotes a posterior probability and the term $w_i(E)$ is the weighting applied by the evidence.

A Revision Rule.—A trivial consequence of the product axiom of probability is a relationship known as Bayes' theorem,

$$p(H_i/E) = \frac{p(E/H_i)p(H_i)}{p(E)} \quad (4)$$

An analogy can be drawn between equations (3) and (4) if a posterior probability is taken to mean the conditional probability $p(H_i/E)$ that the i th hypothesis is true given the evidence, and if the weighting term is identified with the term $p(E/H_i)/p(E)$. There have been recent investigations (ref. 1) to determine whether or not, in his estimation of posterior probabilities, man is a Bayesian (i.e., he applies a revision rule approximating Bayes' theorem which is the formally optimal rule). A dominant finding is that man is conservative: he is inefficient in resolving his uncertainty, as he is unable to make maximum use of the available evidence.

Selection of an algorithm to characterize man's revision rule is complicated not **only** by the question of efficiency, but also by the question of uniqueness. It again seems reasonable to expect that rules for the revision of opinion differ among individuals. Thus, the analogy which has been suggested is certainly but one of many possible. However, we will, nevertheless, accept the analogy for the unique characterization of man's revision making process.

We have assumed there are N hypotheses for every m . Thus the substitution

$$p(E) = \sum_{j=1}^N p(E/H_j) p(H_j) \quad (5)$$

is valid. Making use of the postulated analogy and of equation (5) permits the following definition of the weights :

$$w_i(E) = \frac{p(E/H_i)}{\sum_{j=1}^N p(E/H_j) p(H_j)} \quad (i=1, \dots, N) \quad (6)$$

In this expression the denominator term on the right-hand side can be thought of as a normalization factor which is required in order that the condition,

$$\sum_{i=1}^N p'(H_i) = 1 \quad (7)$$

be satisfied. Therefore, the formal evaluation of the weights $w_i(E)$ can be accomplished once the priors $p(H_i)$ are **known** and the N conditional probabilities $p(E/H_i)$ have been determined. When a prior is revised, the resulting posterior probability becomes the prior for the next revision and so on.

Weighting the Evidence.—If a revision is made, what evidence is used and in what way? An answer to this question depends on the task itself. In the present work we are dealing with a state regulator task in which the human operator actuates a relay to null the output of a second-order dynamic process. At any instant of time during the course of a trial in this task, the signs of the state variables (x, v) and the polarity of the control can be used to distinguish which of four possible situations prevails. Each case is depicted in figure 2 with sketches of segments of the corresponding phase trajectory, and with the controller output called u . A decision to reverse control polarity in each of these four situations presents evidence to the operator which he can use to resolve his uncertainty as to the location of the switch curve.

In the first situation, a decision to switch might result in the outcome illustrated by the first sketch of figure 3. The position at which the trajectory crosses the x -axis is designated as x_k . The theory postulates that the evidence,

$$E_{jk}(x_m): \text{ Switching in the mesh } (x_m, v_j), \text{ when } u \text{ and } x \text{ are} \\ \text{of opposite sign, results in the phase trajectory} \\ \text{crossing the } x\text{-axis between } x_k \text{ and } x_k + \Delta x,$$

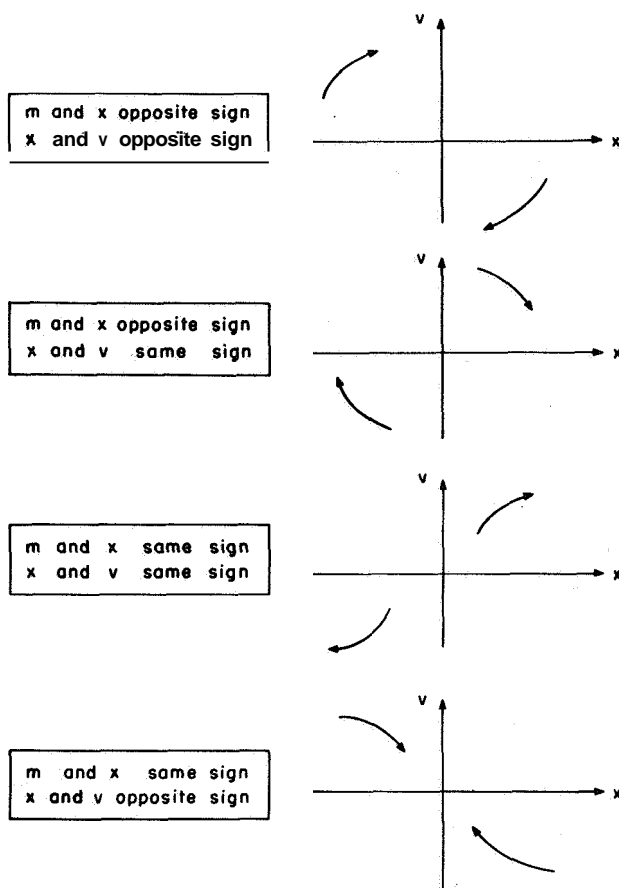


Figure 2.—Four possible cases in state regulator task.

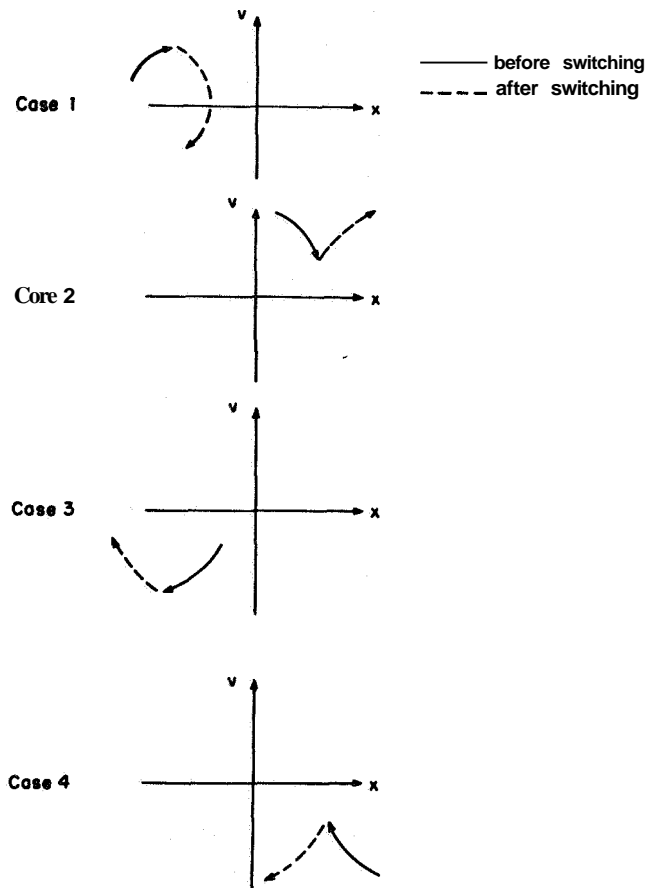


Figure 3.—Results of switching polarity in the four cases.

is used by the operator to test the hypotheses $H_i(x_m)$, where $i=1, \dots, N$. This implies that in order to revise his estimate of $p(H_i)$, the human operator must assign a value to each of the N conditional probabilities, $p(E_{jk}/H_i)$. Collectively, these N conditionals are part of what we call the human operator's "subjective model of the physical world." A subjective model of the physical world summarizes man's beliefs concerning the likelihood of obtaining various outcomes from an experiment when the true state of the world is known. Because they are subjective, models of the physical world will differ among individuals. As was true of revision and selection rules, we again find it expedient to propose a unique characterization. As part of this characterization, we now derive an expression for the conditional probability, $p(E_{jk}/H_i)$.

The Conditionals, $p(E_{jk}/H_i)$.—Assume for the moment that the dynamic process is defined by the differential equation,

$$\ddot{x} = u \quad (8)$$

If the hypothesis H_i x_m that the switch curve passes through the mesh (x_m, v_i) is true and

$1 \leq i \leq n$ (refer to fig. 1), the output of the controller u must have a value somewhere in the range,

$$\frac{(v_i^-)^2}{2x^+} \leq u \leq \frac{(v_i^+)^2}{2x^-} \quad (9)$$

where the plus and minus signs denote the largest and smallest absolute value of the super-scripted state variable in the (x_m, v_i) mesh. For case 1 (figs. 2 and 3), if the control polarity is switched when in mesh (x_m, v_j) and the above hypothesis is true, it is possible for the phase trajectory to cross the x -axis at a point, x_k , somewhere in the interval bounded by

$$d_{\max} = x_m^+ - x_m^- - (v_j^+ / v_i^-)^2 \quad (10)$$

and

$$d_{\min} = x_m^- - x_m^+ + (v_j^- / v_i^+)^2 \quad (11)$$

See the top sketch in figure 4. Equations (10) and (11) are obtained from the first two integrals of equation (8) and the appropriate boundary conditions,

Assume now that the dynamics of the process are not given by equation (8), but by some other differential equation of the second-order, first-degree type we are considering. If the hypothesis $H_i(x_m)$ is true and if in a case 1 situation the control polarity is again switched in the mesh (x_m, v_j) , what values are possible for x_k ? We assert that whatever the dynamics of the unknown process being controlled, we believe it more likely for a crossover to occur within (d_{\min}, d_{\max}) than for it to occur outside this interval. We express this preference by assuming a normal distribution f_{x_k} for the probability that x_k occurs between x and $x+dx$

given $H_i(x_m)$, and taking

$$\bar{x}(m, i, j) = \frac{d_{\max} + d_{\min}}{2} \quad (12)$$

for the mean of the distribution and

$$\sigma_x(m, i, j) = \frac{d_{\max} - d_{\min}}{2} \quad (13)$$

for the standard deviation. See the bottom sketch of figure 4. By definition of a switch curve, an undershoot $x_k < 0$ is not possible for the example portrayed in this figure. Therefore, a truncated distribution is necessary. However, since the weighting of the priors is not seriously affected by ignoring this detail, the truncation is not performed, and the

distribution, therefore, is always defined on $(-\infty, \infty)$. The conditional probability, $p(E_{jk}/H_i)$, is obtained by integrating f_{x_k} between the appropriate limits, that is,

$$\int_{x_k}^{x_k + \Delta x} f_{x_k}(x) dx = p(E_{jk}/H_i) \quad (14)$$

It has been assumed, so far, that $1 \leq i < n$. When $i=n$, $v_1 = 0$ and equation (10) is no longer valid. In this limiting case, $p(E_{jk}/H_n)$ approaches zero, because the magnitude of d_{\max} approaches infinity.

What this means is simply that we are allowing for the possibility, however unlikely, of a switch curve coincident with the x -

axis. When $n < i \leq N$, the hypothesis, $H_i(x_m)$, allows for the possibility that the switch curve lies in the first or third quadrants of phase space, i.e., where x and v have the same sign. For the phase space we have defined ($v = \dot{x}$), the switch curve can not lie in these quadrants, since this would imply $v \neq \dot{x}$. It is hard to imagine what physical reasoning, if any, would lead a human operator to make such a hypothesis. But whatever their rationale, some of them behave (i.e., they switch in these quadrants) as though they temporarily held this belief. Since, by equation (5), we must sum over

all hypotheses, we are compelled to assign values to the conditionals, $p(E_{jk}/H_i)$ where $i=n+1, \dots, N$, even though there is no physical basis for such an assignment. Our judgment of how to make this assignment in some logically consistent manner is to let

$$p(E_{jk}/H_i) = p(E_{jk}/H_n) \quad (i=n+1, \dots, N) \quad (15)$$

At this juncture, we should point out that, as part of man's subjective model of the physical world, the conditional probabilities just introduced represent a conceptualization of momentum. In other words, when these conditionals are used to revise the prior probabilities in a case 1 situation, the posterior probabilities will reflect the belief that if an object is moving with speed $|v_j|$ stops in a distance $|x_k - x_m|$ after switching control polarity, it can be made to stop in a shorter (longer) distance by switching polarity when the speed is less (more) than $|v_j|$. We are assuming, therefore, that the human operator comes to the task with this

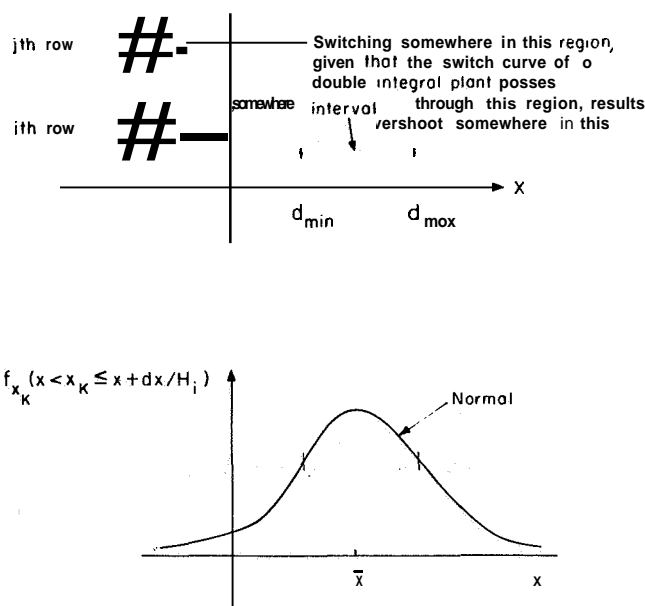


Figure 4.—Weighting of case 1 evidence.

belief and that the distributions of the probabilities $p(E_{jk}/H_1)$ which we have just derived are a suitable characterization of how this belief is conceptualized in the decision center of the human mind. In passing, we note that if the human operator comes to the task with the prior probabilities set to zero in the first and third quadrants, that is, he does not think it probable that the switch curve lies in these regions, he can learn how to control any dynamic process of the class considered simply on the evidence of where the phase trajectory crosses the x-axis after the control polarity is switched in a case 1 situation, provided the evidence is used as we have indicated. Even when the priors are not zero in these quadrants, it is still true that the conditionals, $p(E_{jk}/H_i)$, will enable a subject to resolve his uncertainty as to the location of the switch curve. However, situations such as cases 2 and 3 do provide some additional information which the human operator can use to expedite this resolution and which we now discuss.

Cases 2 and 3.—In a situation similar to the one illustrated by case 2 in figures 2 and 3, reversing the polarity of the control will, in general, cause the phase trajectory to "open up." It is assumed that the sensor can perceive such a pattern; in other words, the human operator recognizes an improper control action. If this is true, it is reasonable to expect that he can also conclude that if the speed, v_j , at the time of switching had been greater, the trajectory would have opened even more than it did and therefore, the hypotheses, $H_i(x_m)$, for $j < i \leq N$ are incorrect and should be rejected. If this evidence is called E_j it can be weighted by revising the priors with the conditional probabilities,

$$p(E_j/H_i) = 0 \quad (i = j+1, \dots, N; j > n) \quad (16)$$

As for the remaining hypotheses, we assume that the human operator does not change their relative ranking on the basis of this evidence, that is,

$$p(E_j/H_i) = 1 \quad (i = 1, \dots, j; j > n) \quad (17)$$

Case 3 is similar to case 2 except that the operator now recognizes he had been using the wrong control polarity, because when he switches, the phase trajectory "closes up." For weighting the evidence, we propose the same conditional probabilities as those given by equations (16) and (17).

The Final Case.—If the operator switches in a case 4 situation, he loses the opportunity to obtain the evidence E_{jk} necessary to evaluate his last switching decision. Persistence in this behavior lengthens the learning time, since the evidence, E_{jk} , is essential to revision making. How the human operator learns that he should wait for a crossover after switching in case 1 will be described using a reinforcement model of the form,

$$p_n = 1 - \alpha(1 - p_{n-1}) \quad (18)$$

where p_n is the probability that the human operator will wait for a crossover after failing to wait n times and α is a learning-rate parameter which determines the strength of each reinforcement.

This completes our discussion of the evidence which is available to the human operator manually controlling a dynamic process. We have postulated how the human operator weights this evidence in resolving his uncertainty as to the location of the switch curve in phase space.

THE EFFECTOR

Response Time.—Executing a response, in the manual control problem we are considering, is a simple task for the human operator's motor system. All that the effector (a finger in this case) must do is depress or release a key which actuates the relay controller. In the present work, it is not essential to provide a description of the time history of the limb movement, since it is only the time the switch actually occurs which matters in our explanation of human learning behavior. We consider the response time to be a uniformly distributed random variable.

Decision Time.—In the decision center, a selection process and a revision process take place. On the basis of the experimental evidence, we have inferred that the times for revision and selection are random variables statistically independent of each other and of the response time. We call the sum of the selection time and the revision time, the decision time. If no revisions are made during a decision cycle, the decision time is determined by the selection time only. The probability density for each component time of decision time is assumed uniform.

THE MODEL.—Our model of human learning behavior in a manual control task is a digital computer program which produces a machine language translation of the theory presented. As Newell, Shaw and Simon (ref. 2) have expressed it, "An explanation of an observed behavior of the organism is provided by a program of primitive information processes that generates this behavior." Herein, these primitive information processes are the selection process, the revision process, etc., which have been set forth by the theory as elements of man's technique for the identification of an unknown control policy. Since the theory has also postulated the rules for combining these processes, the computer program can be written once some final details of the theory have been considered.

For one thing, we have not yet indicated in what order revisions and selections take place. How the human operator establishes priorities in attending to several matters requiring his immediate attention is a difficult question. The order may not be fixed, and it is quite possible that the decision center can interrupt, for instance, the revision process, store the unfinished computations, and attend to a response. Other combinations can also be conjectured. In the model, we assume that revisions come first, selections second, and no interruptions of either are permitted.

For another, we have not specified how many decision cycles are required for the human operator to identify the pattern in the phase trajectory used for evidence in cases 2 and 3. Pattern recognition capabilities vary from one individual to another, and so the number of cycles is not fixed. In the model, we assume that the human operator is capable of detecting the pattern within one decision cycle after the switch occurs.

Finally, we must provide some "numbers" for the parameters which have been left free in the theoretical development. A specification of these parameters corresponds to a specification of the psycho-physiological characteristics of some human operator. As the behavior of the model is governed by the set of numbers chosen, it should be possible to match individual programs with individual human operators. What we mean by "matching" and how this has been accomplished are now discussed.

We have conducted a parametric study of the model on a digital computer. From these results we first found out how these parameters influence the behavior of the model. Then we inferred sets of parameters to provide what we believe to be a representative sample of human operator behavioral simulations. Next we performed a motor-skill experiment and made measurements on the response behavior of human operators. The two samples were then compared statistically to determine whether or not they came from the same parent population,

i. e., whether or not they matched. Figure 5 is a flow chart illustrating the logical steps performed by the computer program. The source program itself may be found in reference 3.

EXPERIMENT

A description follows of a psychomotor experiment performed at M. I. T.'s Man Vehicle Laboratory. Over a 4-month period 50 paid subjects were given the opportunity to learn a manual control task. They were briefed on the task and familiarized with the apparatus, but were not allowed to practice prior to the first trial.

The subject, by actuating a two-position switch, is required to null the initial misalignment between two line segments displayed on an oscilloscope in front of him (see fig. 6). One segment, the left, remains stationary and the displacement of the other relative to it, x , satisfies the differential equation,

$$\ddot{x} = u$$

where u is the switch output and may either be $+U$ or $-U$. As the switch has no off position, once the segment is aligned, rapid polarity changes may be used to simulate an off position and thereby to maintain close alignment. Each subject is given fifty 5-second trials spaced 10 seconds apart. Subject performance on each trial is measured by computing the integral of the absolute value of x over the 5 seconds. This score is reported to the subject immediately after each trial. Every trial starts with the same initial conditions.

RESULTS

THEORETICAL RESULTS.—We have conducted a parametric study of the behavior of the model (i. e., the computer program) to establish how behavior is altered by changes in the psycho-physiological parameters of the model and if the alternations are consistent with our intuitive idea of what should happen. We found that, on the average, the performance of the program in controlling the dynamic process deteriorates whenever the following occur:

- (1) The sensor perceives the state of the dynamic process with greater uncertainty, that is, the mesh size is increased
- (2) The decision center is initially more uncertain of the control policy, that is, the priors are, for example, distributed uniformly or are nonzero in the first and third quadrants
- (3) The decision center requires more time to process information, that is, decision time is increased
- (4) The decision center is slow to recognize that it must wait on the outcome of a response in order to access whether or not it selected the correct choice, that is, α is increased
- (5) The effector requires more time to execute a response, that is, response time is increased.

These findings are consistent with the behavior one would expect to observe in the response performance of any information processing system, these expectations being based, in part, on the predictions of conventional control systems theory. Not to be overlooked, either, is the fact that the program does learn how to control a dynamic process and that the learning process is convergent in all cases. Closed-loop performance of the system, when it is learned, is nearly optimal.

A sequence of five sketches, presented collectively as figure 7, provide a most striking portrayal of learning. One can witness in this sequence the program's progress in resolving its uncertainty as to the location of the switch curve. Each sketch shows a surface, the height of which above the reference plane at the coordinates (x_m, v_i) represents the posterior

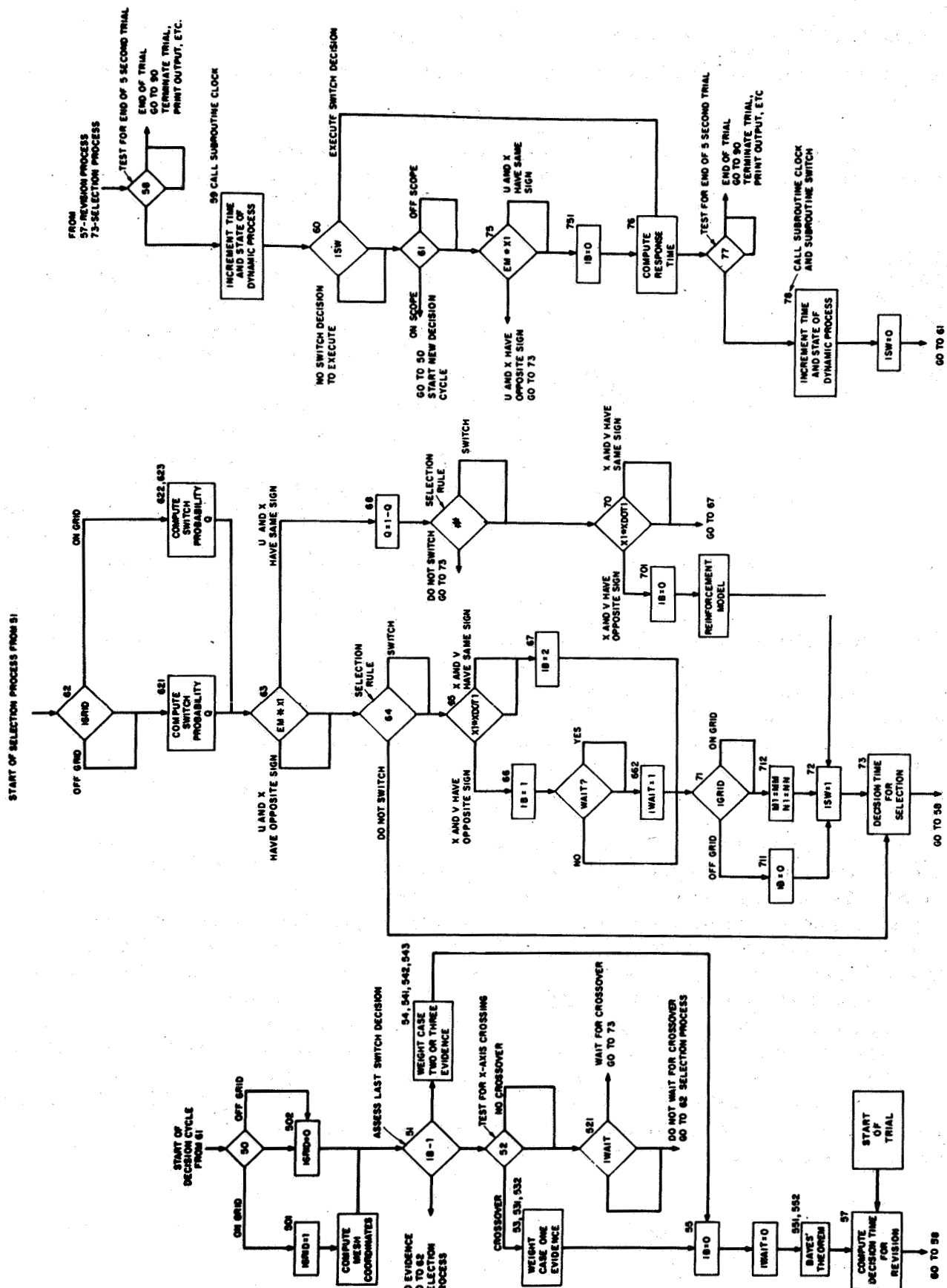


Figure 5.—Flow diagram of source program.

probability $p' [H_i(x_m)]$ at the end of the indicated trial. To give a clearer visual impression, each surface, which is actually formed by a finite set of points, has been filled in and smoothed over.

EXPERIMENTAL RESULTS.—A complete picture of human-operator learning behavior in the psychomotor experiment discussed earlier can be developed from the measurements

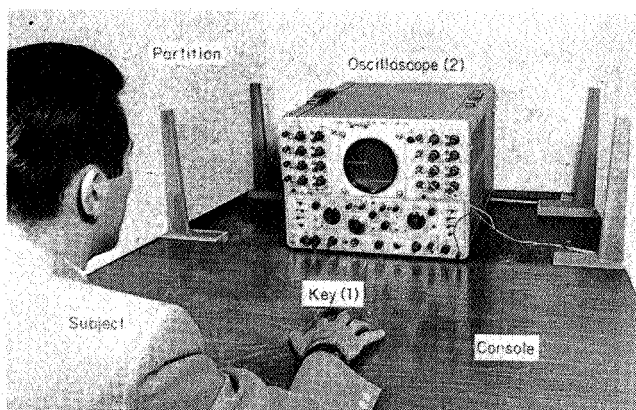


Figure 6.—Equipment for psychomotor experiment,

which were taken of the intervals between successive switches in control polarity. This interval will be referred to as an interresponse time (IRT). From the interresponse time data, the state, (x, v) , of the dynamic process at each switch time has been calculated. To provide a portrayal of learning comparable to figure 7, which depicts the program resolving its uncertainty, we have, for the human operator, taken the statistics on the state variables and computed the ellipsoids of concentration for the first six responses of trials 1, 2, 10, 20, and 50. These appear as a sequence of computer-drawn sketches, collectively called figure 8. An ellipsoid of concentration bounds a two-dimensional region over

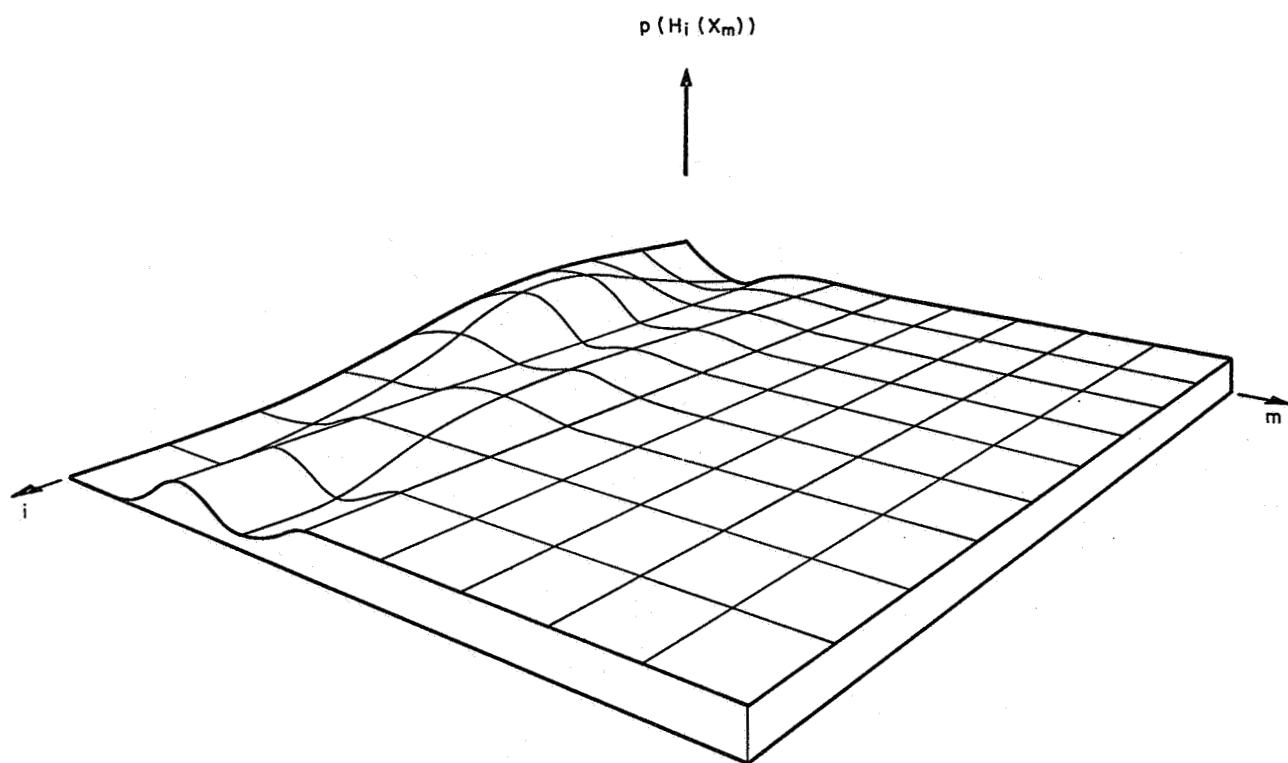
which probability is distributed uniformly such that the first- and second-order moments of the uniform distribution are the same as the actual distribution (see ref. 4, pp. 283-285). A liberal interpretation of this definition, in our case, is to say that the n th region shows where, in state space, "most" subjects made the n th response. The shrinking and reorientation of the ellipses are a vivid illustration of the ensemble's progress in identifying a control policy. In addition to this portrayal of learning, average transient responses for some of the same trials have been calculated from the state data and are presented collectively as figure 9.

THEORY AND EXPERIMENT COMPARED.—From the figures presented in the previous sections, one can easily develop a qualitative appreciation for the individual differences exhibited in the learning behavior of both subjects and programs. For example, initial score, level of asymptotic performance, and rate of score change are some of the readily discernible indicators provided by the integrated squared error which are useful in comparing the motor-skill behavior of the programs with that of the subjects. What is important now is to answer the question of whether or not the theory developed herein is a credible explanation of human learning behavior, particularly of intersubject, intrasubject variability. For this purpose, it is desirable to make a comparison of the subject ensemble and the test sample of programs on some quantitative basis. This has been done and is discussed next.

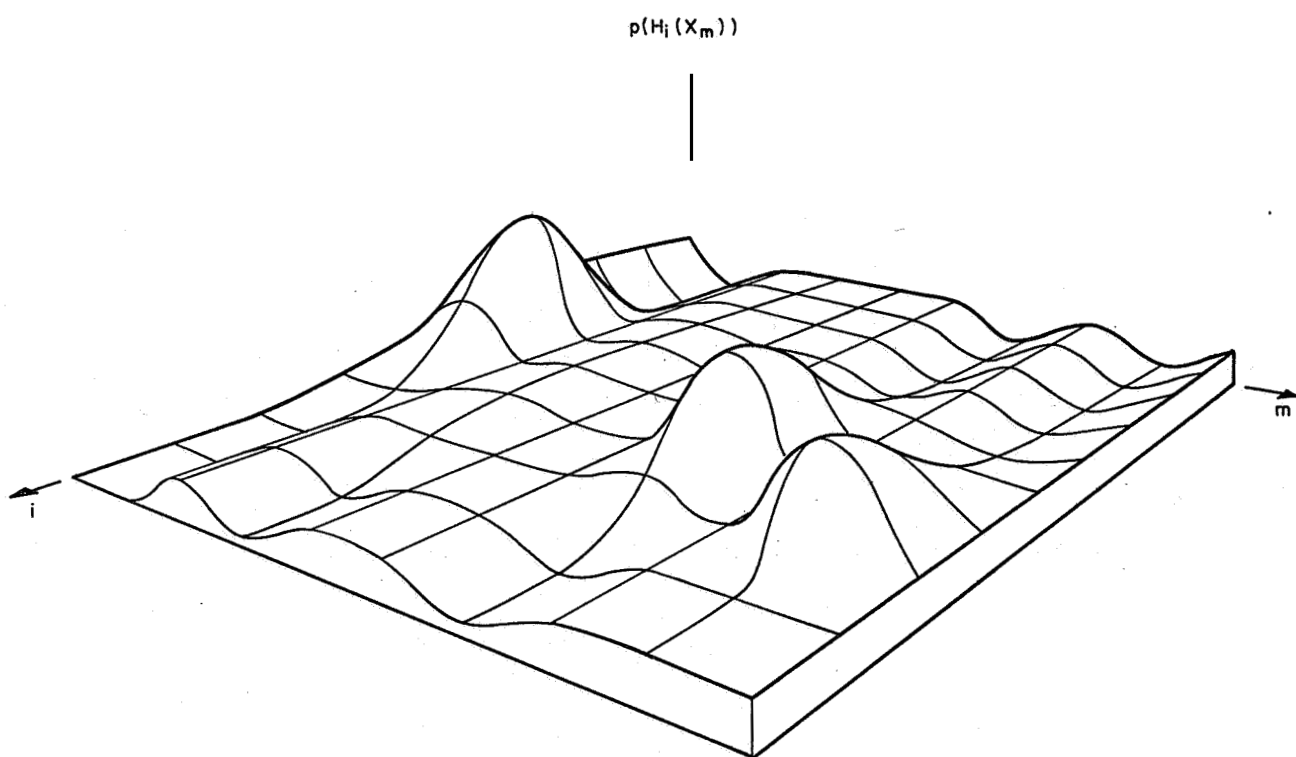
To establish the "similarity" between the behavior of the fifty subjects and ten programs of a test sample, the Mann-Whitney U test was applied to each of the first four interresponse times of each twelve trials. On a given trial for a given response, the sample of subject

IRT's $(IRT_i \text{ where } i=1, 2, \dots, n_1)$ and the sample of model IRT's $(IRT_j \text{ where } j=1, 2, \dots, n_2)$ are arranged in order; the statistic, U , counts the number of times a member of the first sample exceeds a member of the second sample. If

$$z = \left| \frac{(U - \bar{U})}{\sigma_U} \right| \quad (19)$$

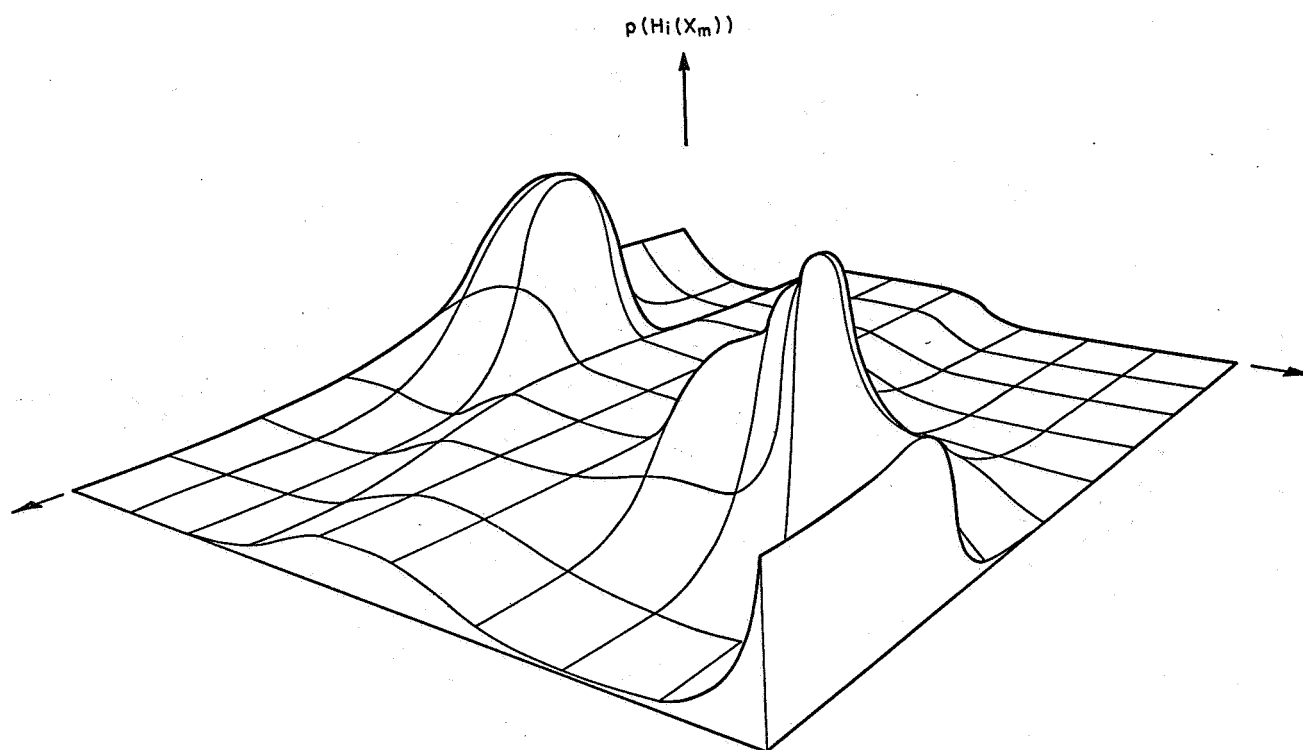


TRIAL 1

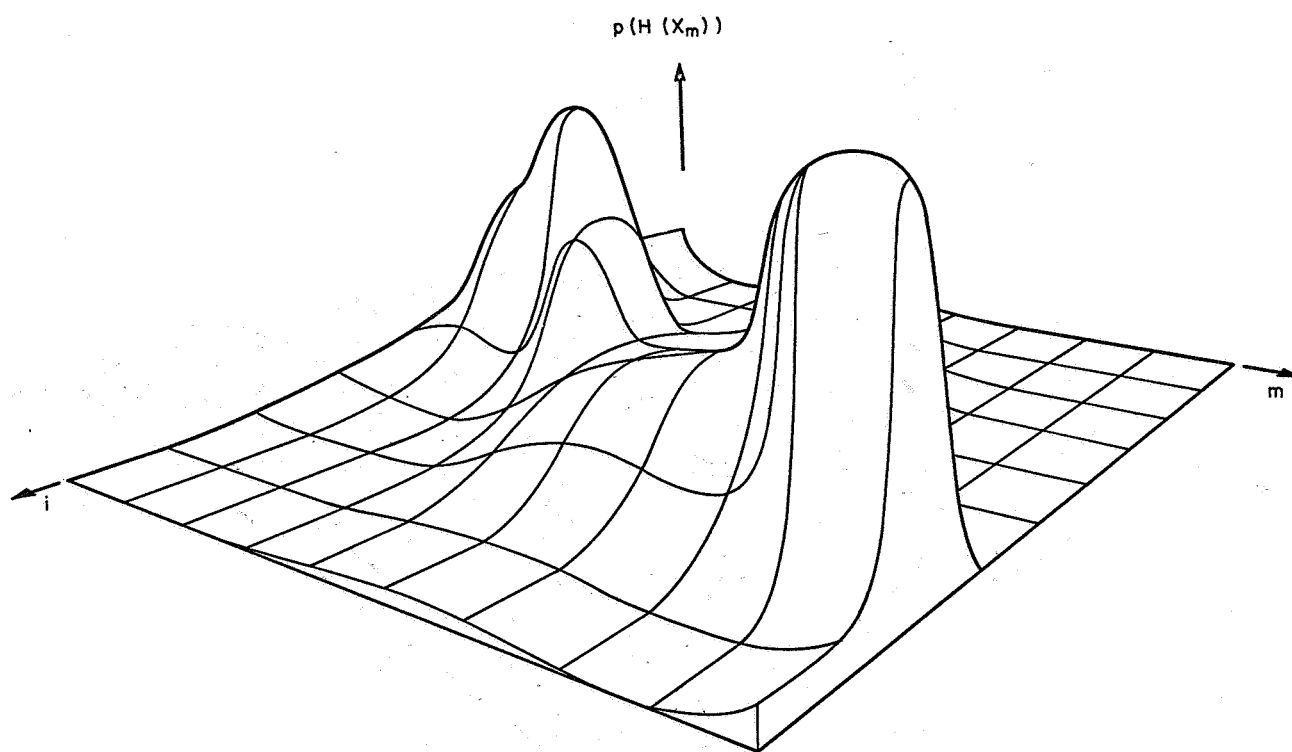


TRIAL 2

Figure 7.—Model learning behavior.

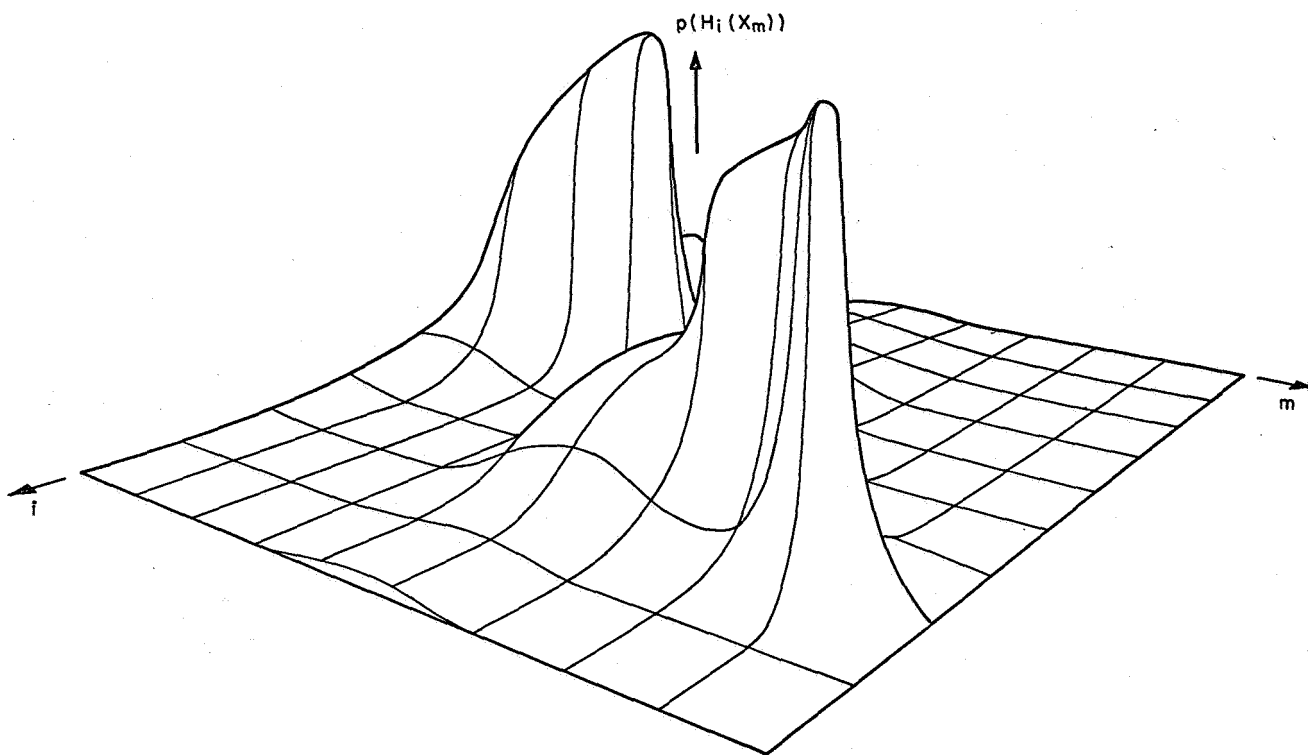


TRIAL 5



TRIAL 10

Figure 7.—Continued.



TRIAL 50

Figure 7.—Concluded.

is greater than 2.58 under the null hypothesis ($f=g$), the test is considered significant at the 1 percent level, and the hypothesis of identical distributions is rejected. Table 1 lists the values of z calculated for the 48 test cases.

At the 1-percent level, the table shows that 81 percent of the cases pass the test, that is, the hypothesis of identical distributions is acceptable. Those cases which fail the test are confined to the third and fourth responses on trials after the fifth. The reason for these failures is that subjects develop an open-loop technique for responding when the dynamic process' state is close to the origin. This mode of behavior is an attempt by a subject to simulate an off position (as he was instructed) with the controller by rapidly alternating control polarity. In this mode, the subject effectively ignores state information until such a time as the error exceeds some tolerance level, and then he reverts back to a closed-loop mode of responding. Clearly, the theory does not account for this, since the program makes but one response per decision cycle and does not set off preprogrammed sequences of responses. Aside from this discrepancy, the results of the "U" tests are quite favorable and offer no cause to reject the hypothesized identity of the two population distributions. Note: at the 1-percent level of significance, the probability of obtaining a Z -value greater than 2.58 when comparing two samples is, by definition, 0.01, given that the hypothesis $f=g$ is true.

We have also applied the Mann-Whitney U test to samples of the integrated squared error scores on the same trials as before. The results, which are presented in table 2, show that only one case is significant at either the 1-percent or 5-percent ($Z \geq 1.96$) level: For the human operator sample, we selected the first 10 subjects instead of using the entire ensemble. Performing this test on the scores is a less sensitive measure of the credibility of our theory than performing it on the IRT's, since the integration to obtain a score masks

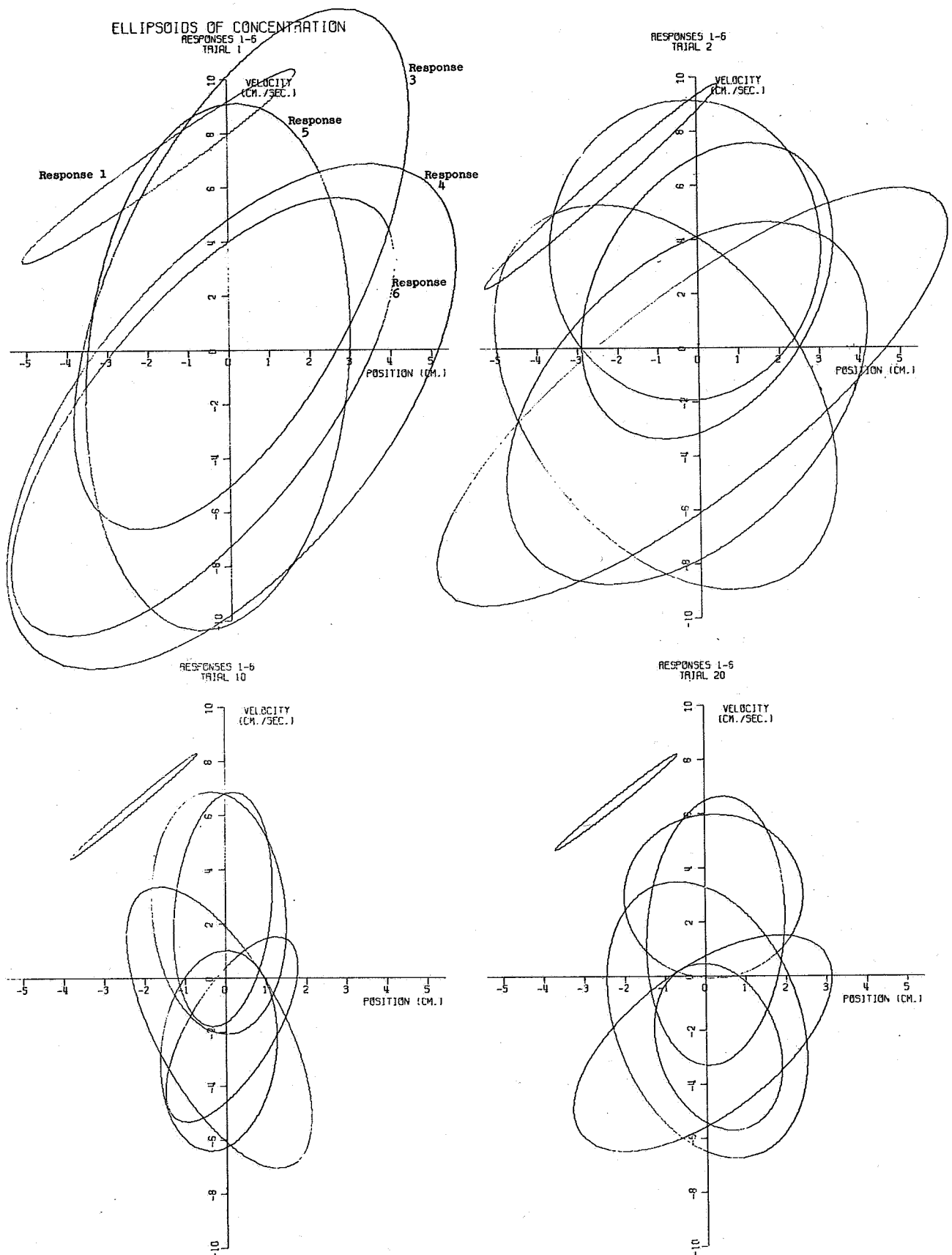


Figure 8.—Ellipsoids of concentration.

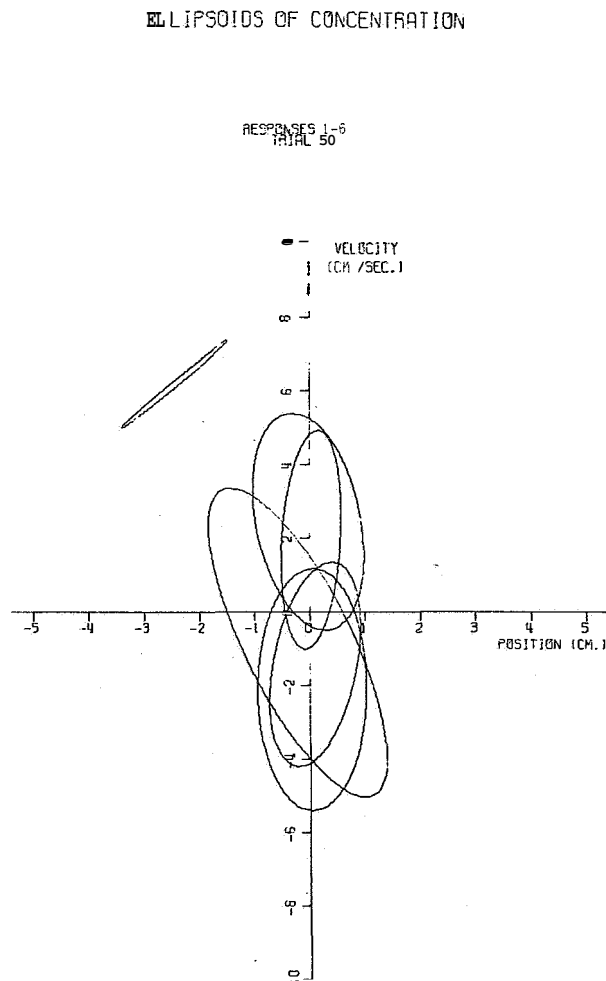


Figure 8.—Concluded. Ellipsoids of concentration.

the detailed structure of the response behavior, and, therefore, discrepancies in this structure can be obscured from detection. Testing the IRT's, on the other hand, subjects the finest grain measurement we have available on the response behavior to the scrutiny of a powerful nonparametric statistical test.

DISCUSSION OF RESULTS.—In the preceding section, we made a statistical comparison of a sample of human operator behavioral simulations obtained from computer program executions and a sample of human operator behavioral data obtained from a psychomotor experiment. Through this comparison we have sought to determine whether or not the samples came from the same parent population, that is, whether they are statistical images of one another. The results of the comparison show that, with the exceptions noted, there is no statistical reason for rejecting the hypothesis of identical parent population distributions. Although this is a favorable outcome and offers us a quantitative basis for having confidence in the proposed theory, we hesitate to conclude that this result, by itself, is sufficient evidence upon which to argue for the credibility of the theory. We hesitate because of the inherent limitations of any statistical test, namely, the

possibility that a false hypothesis can be accepted and the possibility that other theories can pass the same test. However, if this result is weighed together with the results of the parametric study and the experimental findings referenced in support of assumptions made in the theoretical development, the case for credibility is strengthened considerably; therefore, there is a rational basis for accepting the theory's explanation of human learning behavior in the type of manual-control task considered.

CONCLUSION

SUMMARY.—We have developed a theory for the explanation of human learning behavior in a manual-control task. In explaining how the human operator acquires a motor skill, we have endeavored to account for the intersubject, intrasubject variability which is observable in psychomotor experiments. This variability has been attributed to the stochastic nature of human information processing, which we have assumed to be a sequential operation involving three subsystems: the sensor, the decision center, and the effector. Each of these components has been treated as a probabilistic system, and stochastic descriptions of how they function have been provided. Our interpretation of Bayesian statistics for the characterization

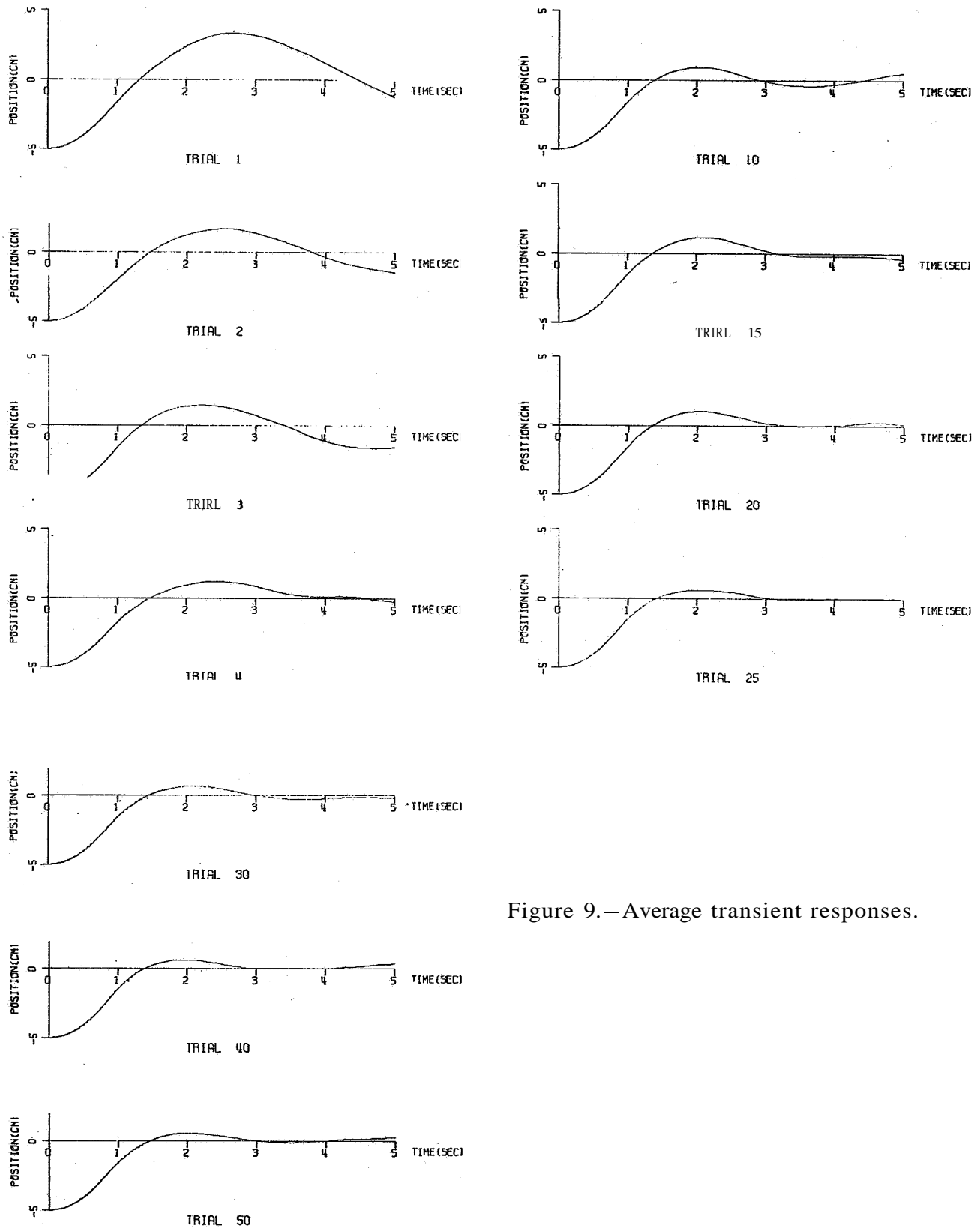


Figure 9.—Average transient responses.

of the decision center's decision making has been, perhaps, our most important contribution to the understanding and conceptualization of human learning behavior. From the theory we have derived a model of human learning behavior in a manual-control task. This has been

TABLE 1.—z-VALUES FOR
MANN-WHITNEY U TEST OF
INTERRESPONSE TIMES

| Trial | Response | | | |
|-------|-------------|-------------|------|------------|
| | 1 | 2 | 3 | 4 |
| 1 | 1.74 | 0.77 | 2.17 | 0.96 |
| 2 | .63 | 1.11 | 2.16 | .07 |
| 3 | 1.92 | .14 | 1.25 | .74 |
| 4 | .95 | .52 | 1.70 | .40 |
| 5 | 1.48 | 1.33 | .89 | .43 |
| 6 | 1.86 | .50 | 3.34 | 3.06 |
| 7 | .73 | 1.07 | 3.17 | 2.48 |
| 8 | 1.11 | .89 | 2.34 | 1.88 |
| 9 | 2.12 | .12 | 2.40 | 2.91 |
| 10 | .40 | .85 | 2.87 | 2.80 |
| 25 | .71 | .93 | 3.14 | 3.22 |
| 50 | .75 | 1.62 | 4.10 | 2.50 |

TABLE 2.—z-VALUES
FOR MANN-WHITNEY U
TEST ON INTEGRATED
SQUARED ERRORS

| Trial | z |
|-------|-------|
| 1 | 0.903 |
| 2 | .376 |
| 3 | .827 |
| 4 | .526 |
| 5 | .300 |
| 6 | .376 |
| 7 | .450 |
| 8 | .266 |
| 9 | 2.630 |
| 10 | 1.352 |
| 25 | 1.276 |
| 50 | 1.051 |

accomplished by a translation of the theory into the machine language of a digital computer. A set of read-in parameters, corresponding to human psycho-physiological characteristics, gives the model an individuality. Consequently, we have been able to execute a number of computer programs, which, on the basis of a hypothesis test, have been shown to be a statistical image of an ensemble of human operators. The number of parameters required to establish the identity of the model is relatively small considering the complexity of the process being simulated and the detailed similarity it **offers**.

GENERALIZATIONS.—We now explore the possibility of generalizing the approach of the theory developed herein, for the purpose of explaining human learning behavior in other manual-control task contexts.

Continuous Controller.—The first extension we wish to consider is a task where the controller output can be varied continuously over a bounded range by the operator, but where there is otherwise no difference from the task we have already treated. In the task we have treated, we assumed that in the decision center there are stored probabilities for the $M \times N$ hypotheses,

$$H_i(\mathbf{x}_m) : \text{the switch curve passes through the mesh } (\mathbf{x}_m, \mathbf{v}_i)$$

Let us reword these hypotheses so that they read,

$$H_i(\mathbf{x}_m) : \text{the controller output, } u, \text{ in the mesh, } (\mathbf{x}_m, \mathbf{v}_i), \text{ equals } u_0,$$

where u_0 may either be $+U$ or $-U$. Written in this form, it is clear that the control alternatives are $\pm U$ and that the probability distribution for $H_i(\mathbf{x}_m)$ is discrete. To make a transition to a continuous controller, we write,

$H_i(x_m)$: the controller output u in the mesh (x_m, v_i) equals or is less than u_0 ,

where u_0 is now a continuous variable defined on the interval $(-U, +U)$, and the distribution of u is also continuous. With this definition one can trace through the steps of the derivation and see that basically the only change necessary in the development is to replace summation signs by integrals and discrete distributions by continuous ones, where appropriate.

Pursuit Tasks.—In the state-regulator problem we have considered, the terminal state to which the dynamic process is being forced, is fixed. In a pursuit task, the terminal state may change with time, and so, control decisions must be based on an estimate of the anticipated terminal state at the expected time of convergence. In other words, the decision center must make predictions of the future course of events. Therefore, in the information processing sequence, we must insert a prediction operation. In addition, the center's memory must store not only the probabilities of response alternatives for reaching the null state, but also the probabilities for reaching all other meshes in state space which are possible location of the terminus.

Other Tasks.—Extensions of the theory to other task contexts, including compensatory tracking problems and controlling dynamic processes not in the class to which we have restricted the present development, are also conceivable. However, in such task contexts it is doubtful that our interpretation of the evidence, E , is still applicable. Since we have not, as yet, studied these situations in any detail, we will not speculate on how the theory may be generalized to handle them.

APPLICATIONS

Adaptive Control.—One application of this work, which we would like to discuss, is in the field of adaptive control systems. If the computer program presented in this work is examined carefully, it can be seen that without the input-output statements and the superfluous subroutines for keeping score, simulating the dynamic process, etc., the logic of the program requires relatively few FORTRAN statements. In fact, if the priors are set to 0 in the first and third quadrants, thereby eliminating the need for weighting case 2 and 3 evidence; if p_0 is set to 1, if the sensor function is deleted, and if a few other nonessentials are removed, the program reduces to a very few statements. In such a form, it does not appear that there would be any great difficulty in constructing a special-purpose digital computer to execute the control logic. If provision is made not to allow the probabilities to go to zero, the program will learn, unlearn, and relearn control policies. A modified version of our model of human learning behavior we believe, therefore, has the potential to perform as the logic element of an adaptive control system.

Psychomotor Testing.—A second application, for which the theory holds promise, is in the selection of pilot trainees. The introduction in World War II of psychomotor testing, by the military to select flight crews, marked the beginning of a continuing search for improved techniques to determine the aircrew potential of individual applicants. In the development of our theory, we have identified explicitly the determinants of human behavior in a manual-control task. These included the selection rule, revision rule, prior probabilities, decision and response times, etc. Reflected in these determinants of behavior are the operator's past experience in manual-control tasks, the efficiency of his information processing, his physiological limitations, and the like. While it is true that these qualities alone are not sufficient to judge the aircrew potential of an individual, they are, nevertheless, important aptitude indicators. It may be possible, therefore, to devise a method, based on the theory, for statistically inferring the characteristics of an individual's information processing system from his performance in a single manual-control task. We have, in fact, already done something

similar to this in determining the model parameters for generating our sample of behavioral simulations.

REFERENCES

1. Beach, L. R.: Accuracy and Consistency in the Revision of Subjective Probabilities. Human Factors in Electronics, vol. HFE-7, no. 1, Mar. 1966.
2. Newell, A.; Shaw, J. C.; and Simon, H. A.: Elements of a Theory of Human Problem Solving. Psychol. Rev., vol. 65, no. 3, 1958, pp. 151-166.
3. Preyss, A. E.: A Theory and Model of Human Learning Behavior in a Manual Control Task. Sc. D. Thesis, M.I.T., Feb. 1967.
4. Cramer, H.: Mathematical Models of Statistics. Princeton Univ. Press, 1946.

BIBLIOGRAPHY

- Adams, J. A.: Motor Skills. Ann. Rev. Psychol., vol. 15, 1964, pp. 181-202.
- Bilodeau, E. A.; and Bilodeau, I. McD.: Motor Skills Learning. Ann. Rev. Psychol., vol. 12, 1961, pp. 243-280.
- Good, I. J.: The Estimation of Probabilities. Res. Monograph no. 30, The M. I. T. Press, 1965.
- Hohle, R. H.: Inferred Components of Reaction Times as Functions Foreperiod Duration. J. Exp. Psychol., vol. 69, no. 4, 1965, pp. 382-386.
- Mann, H. B.; Whitney, D. R.: On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other. Annals Math. Stat., vol. 18, 1947, pp. 50-60.
- Welford, A. F.: The Measurement of Sensory-Motor Performances: Survey and Reappraisal of Twelve Years' Progress. Ergonomics, vol. 4, 1960, pp. 189-230.
- Welford, A. T.: The Psychological Refractory Period and the Timing of High-Speed Performance—A Review and A Theory. Brit. J. Psychol., vol. 43, 1952, pp. 2-19.
- Wilde, R. W.; and Westcott, J. H.: The Characteristics of the Human Operator Engaged in a Tracking Task. Automatica, vol. I, 1962, pp. 5-19.
- Young, L. R.; and Stark, L.: Biological Control Systems—A Critical Review and Evaluation. NASA CR-190, 1965.

27. Manual Time-Optimal Control for High-Order Plants

Syozo Yasui and Laurence R. Young
Massachusetts Institute of Technology

N68-15928

The concept of a switching surface is applied to closed-loop, manual, time-optimal, bang-bang control of high-order systems by visually displaying the instantaneous position-velocity error spot and the switching locus simultaneously on the phase plane. Switching locus is a curve describing the intersection between the switching surface and the plane parallel to the phase plane, for example, representing the locus in the instantaneous acceleration plane for third-order plants. The operator's task is to achieve minimum time response by reversing control polarity when the moving state point intersects the moving switching locus.

The general procedure is outlined for n th order plants with real poles, and a technique suitable for an analog computer to generate the exact or approximate switching locus is proposed. The second-order plants, $1/s^2$ and $1/s(s+a)$, are presented as preliminary examples, for which the switching loci do not move. The third-order plant, $1/s^3$, is studied in detail as an example of higher order plants. Satisfactory experimental results have been obtained for all examples given.

The phase plane is both a useful tool for studying human behavior in bang-bang control and a potential aid for human operators involved in difficult manual control systems (refs. 1 and 2). Figure 1 shows a manual control system utilizing the $e-\dot{e}$ display. The human operator's task is to drive the dot to the origin (plant equilibrium) from some initially disturbed state as rapidly as possible. A human operator can accomplish this fairly well for stable second-order plants, including critically stable plants, $1/s^2$ and $1/s(s+\alpha)$, and he requires no more than four times the theoretical minimum time for each trial. Although response times are less satisfactory, a human operator can control critically stable third-order plants by using the "chatter mode" when approaching the origin in the second or fourth quadrant of the $e-\dot{e}$ plane (ref. 3). The human operator is incapable of stabilizing any undamped plant of higher order without additional information.

In an effort to provide additional information in the form of a switching line, time-optimal control theory has been applied to a plant of n th order with real poles and maximum magnitude of control force being m . Under these conditions the theory states (refs. 4 and 5) the following:

- (1) The time optimal control force must be $\pm m$
- (2) The control switching may occur at most $n-1$ times
- (3) The switching criteria are successively determined by the switching hypersurface S_{n-1} and its partitioned subsets, S_{n-2} , S_{n-3} , ..., S_1 . In the $e-\dot{e}$ plane, the switching criterion is the curve representing the intersection of the switching subspaces with another subspace determined by the instantaneous higher derivatives of the plant. This curve is called the switching locus.

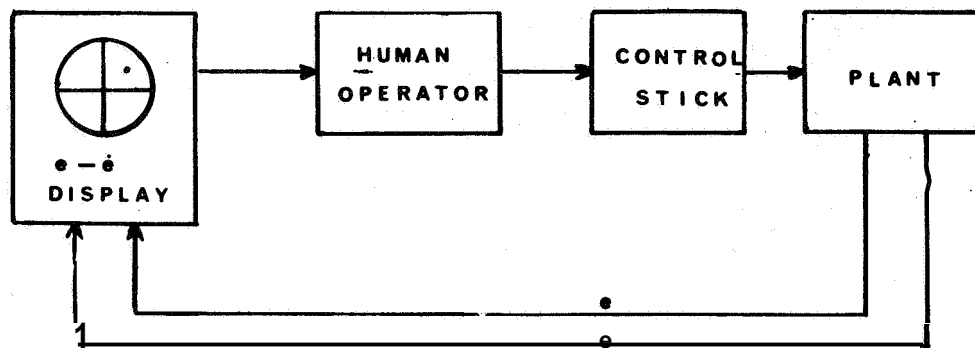


Figure 1.—Manual control system with $e - \dot{e}$ display.

The switching locus moves according to higher error derivatives for greater than second-order plants, dividing the $e - \dot{e}$ plane into a $+m$ control region and a $-m$ control region. The initial choice of control force should be made in such a way that the state point and the switching locus approach each other. A control polarity reversal must take place when they meet. From the moment the n -dimensional state point intersects the switching hypersurface, it describes an optimum trajectory along the hypersurface S_{n-1} until the switching subset S_{n-2} is reached. As soon as S_{n-2} is reached, a second switching must be made. Thereafter, the state point will proceed along the subset S_{n-2} until S_{n-3} is reached. In this manner the state point will reach the subset S_1 , which is an n -dimensional curve passing through the origin, and thus finally arrive at its destination.

The entire state space S_n is divided by S_{n-1} into S_n^+ and S_n^- corresponding to control regions $+m$ and $-m$, respectively. S_{n-1} is again divided by S_{n-2} into S_{n-1}^+ and S_{n-1}^- ; in this manner, each subset is divided by the succeeding subset until the final one, S_0 , is degraded to the origin.

Let us now return to the $e - \dot{e}$ plane and interpret the above statements in terms of the switching locus for $n=3$. If the first switching reversal is at approximately the correct time, the dot (e, \dot{e}) will automatically follow the switching locus until the subset S_{n-2} is reached in the state space. That is, the next switching time might be at the moment when the dot (e, \dot{e}) begins to leave the switching locus. For the third-order plant, $1/s^3$, this second switching criterion is simply given by a corner of the switching locus, since S_1 divides S_2 . After the second switching, the dot (e, \dot{e}) stays at the moving corner point of the switching locus, eventually reaching the origin.

GENERAL DESCRIPTION OF SYSTEM REALIZATION

Figure 2 is a block diagram for a manual time optimal control system for n th order plants. The switching hypersurface S_{n-1} can be expressed as

$$S_{n-1}(x_1, x_2, \dots, x_n) = 0 \quad (1)$$

where x_1, x_2, \dots, x_n represent $e, \dot{e}, \dots, e^{(n)}$, respectively. Note that all subsets $S_{n-2}, S_{n-3}, \dots, S_1$ are contained in this equation, since $S_{n-1} \supset S_{n-2} \supset \dots \supset S_1$. The switching-locus generator measures the instantaneous higher derivative outputs x_3, x_4, \dots, x_n , substitutes the measured values in equation (1), and solves for a locus of x_1 and x_2 .

An exact closed-form expression of this type is difficult to obtain for higher order cases. The expressions for S_{n-1} for third-order plants which have been investigated are too

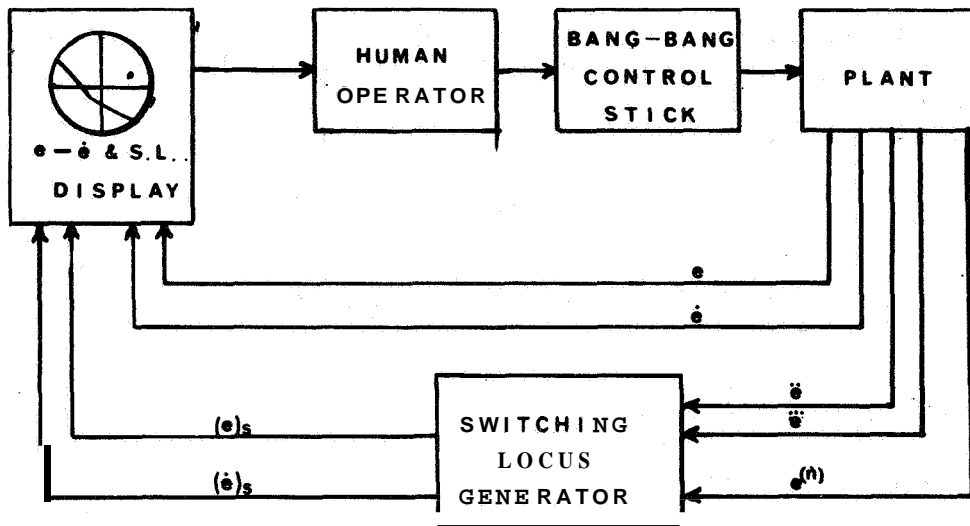


Figure 2.—Proposed manual time optimum control system using optimum switching locus.

complicated to be applied in this system (refs. 6 and 7), so a set of parametric equations derived by the reversed time technique (ref. 4) is used to express $S, -\sim$.

Let us solve the differential equation characterizing the plant dynamics backward in time τ , with null initial conditions and the control-force backward time history $u(\tau)$ such that

$$u(\tau) = \begin{cases} i m & 0 \leq \tau \leq \tau_1 \\ \pm m & \tau_1 \leq \tau \leq \tau_1 + \tau_2 \\ \vdots & \vdots \\ \pm (-1)^{n-1} m & \tau_1 + \tau_2 + \tau_{n-2} \leq \tau \leq \tau_1 + \tau_2 + \dots + \tau_{n-1} \end{cases}$$

Then, the solution will be expressed as

$$\left. \begin{aligned} x_1 &= x_1(\tau_1, \tau_2, \dots, \tau_{n-1}, \pm m) \\ x_2 &= x_2(\tau_1, \tau_2, \dots, \tau_{n-1}, \pm m) \\ &\dots \dots \dots \\ x_n &= x_n(\tau_1, \tau_2, \dots, \tau_{n-1}, \pm m) \end{aligned} \right\} \quad (2)$$

This set of equations is the parametric representation of the switching hypersurface S_{n-1} . Note that substituting $\tau_{n-1}=0$ in equation (2) yields the subset S_{n-2} ; substituting $\tau_{n-1}=\tau_{n-2}=0$ yields S_{n-3} , and so on.

Eliminating all τ 's except one, τ_k , from equation (2) yields

$$\left. \begin{aligned} x_1 &= x_1(x_3, x_4, \dots, x_n, \pm m, \tau_k) \\ x_2 &= x_2(x_3, x_4, \dots, x_n, \pm m, \tau_k) \end{aligned} \right\} \quad (3)$$

The choice of τ_k depends on x_3, x_4, \dots, x_n and the control polarity, and each τ_k will define a part of the switching locus. Further discussion of the above situation appears in the section on the plant $1/s^3$.

The set of parametric equations (3) is the switching locus generating equation suitable for an analog computer capable of repetitive operation, since τ_k may be treated as an independent time variable running from zero to a certain value repeatedly. Furthermore, each function of equations (3) might represent a solution to some differential equation. If this is the case, the switching locus can be more conveniently generated by repetitive solutions to such differential equations with initial conditions being functions of x_3, x_4, \dots, x_n . In fact, simple differential equations can be obtained for plants without any time constant such as $1/s^2$ and $1/s^3$, since equations (2) and (3) for such plants are polynomial functions.

For plants involving some time constants, on the other hand, equations (2) will contain some exponential functions, and therefore equations (3) will also involve certain exponential functions and perhaps a logarithmic function as well. For this reason, the appropriate simple differential equations may not be obtainable. However, let us assume that the time constant is large enough to permit

$$(\text{Time to go})/(\text{Time constant}) \ll 1$$

or that we are only interested in a region in which the origin can be reached in much less than one time constant. Then equations (2) and (3) can be approximated by polynomial functions by taking finite Maclaurin series expansions for the exponential functions and/or the logarithmic functions. The appropriate differential equations for generating the approximate switching locus could be thus obtained. Figure 3 shows the experimental site for the test system.

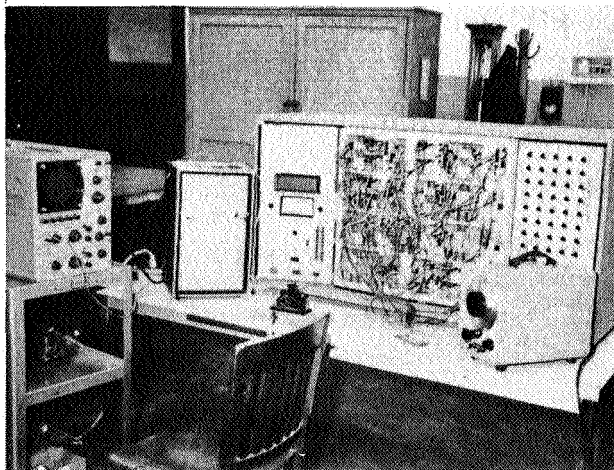


Figure 3.—Photograph of experimental site.

EXAMPLES

EXAMPLE 1: $\frac{1}{s^2}$

The differential equation for the plant dynamics is

$$\frac{d^2 e}{dt^2} = \pm m$$

The backward equation then becomes

$$\frac{d^2 e}{d\tau^2} = \pm m$$

It is easy to conclude that the backward equation directly leads to the switching-locus generating equation with null initial conditions for second-order plants. Substituting $(e)_s = (x)_s = x$ and $\frac{d(e)}{dt} = -(x_2)_s = -y$, we can decompose the backward equation:

$$\frac{d^2 x}{d\tau^2} = \pm m$$

$$\frac{dx}{d\tau} = -y$$

with null initial conditions. It can be seen that the solution

$$x = \pm \frac{m}{2} \tau^2$$

$$y = \pm m\tau$$

leads to the familiar form of the switching locus

$$x = -\frac{1}{2}|y|y$$

Figure 4 shows the analog computer realization of this solution. Figure 5 shows the results of a typical human operator's trial with no switching locus for $1/s^2$. The initial condition was $(e=2.0, \dot{e}=0.5 \text{ sec}^{-1})$ and $m=0.125 \text{ sec}^{-2}$. Figure 6 shows the results when the exact switching locus display was added. The control reversal occurred once, precisely as predicted by the theory, and the time consumed was reduced by approximately 60 percent. Figure 7 shows the manually controlled, time-optimum trajectory and the exact switching locus on the $e-\dot{e}$ display.

EXAMPLE 2: $\frac{1}{s(s+\alpha)}$

As in example 1, the generating equations are obtained

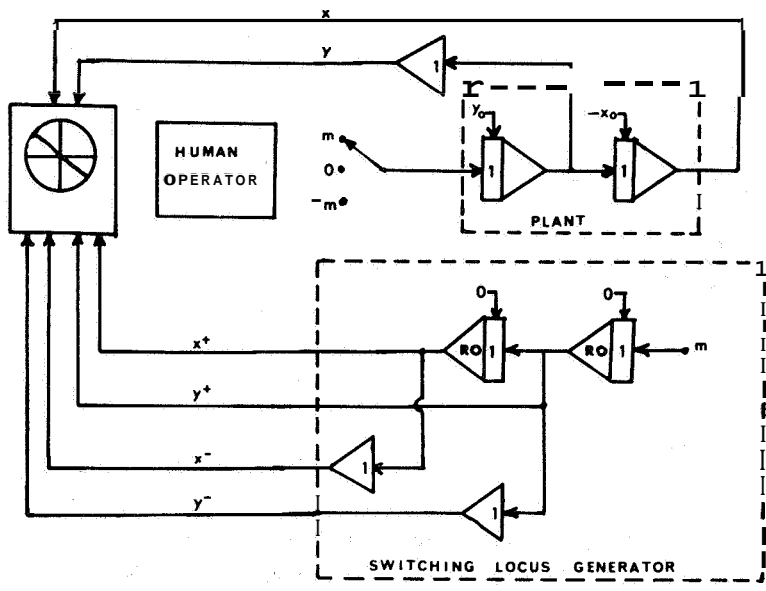


Figure 4.—Analog computer realization of manual time optimum control for $1/s^2$. Integrators labeled RO perform repetitive operation.

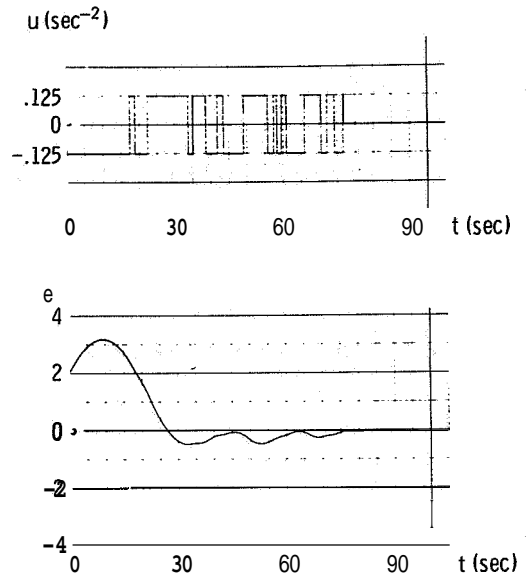


Figure 5.—A typical control force u and error e versus time obtained by experienced human operator stabilizing $1/s^2$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$), when state dot (e, \dot{e}) is displayed but no switching locus is displayed.

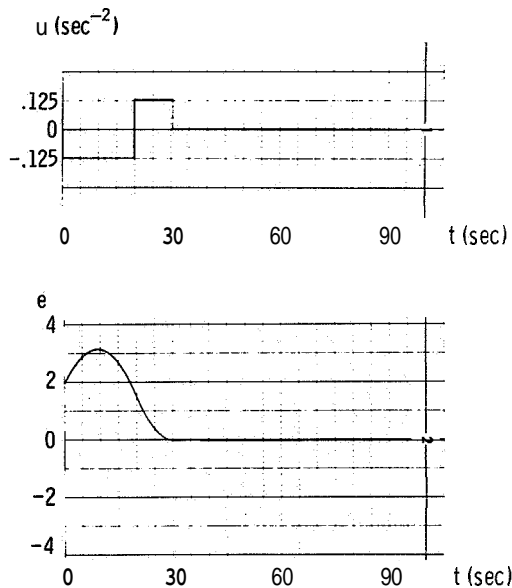


Figure 6.—A manual time optimum control result for $1/s^2$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$), when both state dot (e, \dot{e}) and exact optimum switching locus are displayed.

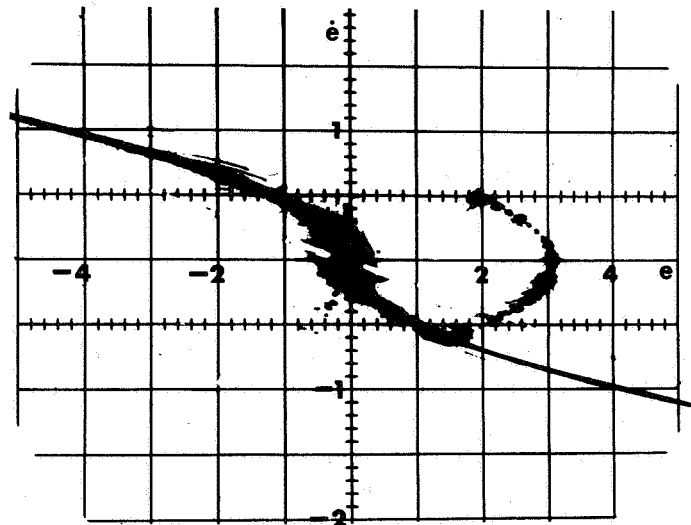


Figure 7.—Photograph of a time optimum trajectory and exact switching locus in e - \dot{e} display for $1/s^2$ with initial conditions ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$).

$$\frac{d^2x}{d\tau^2} - \alpha \frac{dx}{d\tau} = \pm m$$

$$\frac{dx}{d\tau} = -y$$

with null initial conditions. Figure 8 shows the analog computer realization. Figure 9 shows an experienced operator's response for the plant $1/s(s+0.16)$ with no switching locus. The initial conditions were the same as in example 1. Figure 10 shows the results when the exact switching locus was displayed in addition to the state dot. The optimum response was again achieved manually by just one control reversal, and the time consumed was reduced by approximately one-half. Figure 11 shows the above results on the e-ê display. In spite of the presence of time constants, the exact generating differential equations were easily obtained, but only because the plant is of second order. It may be instructive to develop the approximate generating equations according to our earlier suggestion.

The exact generating equations are

$$x = \pm \frac{m}{\alpha^2} \left(e^{\alpha\tau} - \alpha\tau - 1 \right)$$

$$y = \pm \frac{m}{\alpha} \left(e^{\alpha\tau} - 1 \right)$$

Expanding $e^{\alpha\tau}$ in a Maclaurin series gives

$$e^{\alpha\tau} = 1 + \alpha\tau + \frac{(\alpha\tau)^2}{2} + \frac{(\alpha\tau)^3}{6} + \frac{(\alpha\tau)^4}{24} + \dots$$

If we assume $\alpha\tau \ll 1$ and take the first five terms of the series, the resulting approximate expression is

$$x = \pm m \left(\frac{\tau^2}{2} + \frac{\alpha}{6} \tau^3 + \frac{\alpha^2}{24} \tau^4 \right)$$

Note that substituting $\alpha=0$ in the above equations yields the exact locus-generating equations for the plant $1/s^2$. The differential equations for the approximate Switching locus are such that

$$\frac{d^4x}{d\tau^4} = \pm m\alpha^2$$

$$y = -\frac{dx}{d\tau}$$

with initial conditions:

$$x(0)=0, \dot{x}(0)=0, \ddot{x}(0)=\pm m, \ddot{\ddot{x}}=\pm m\alpha$$

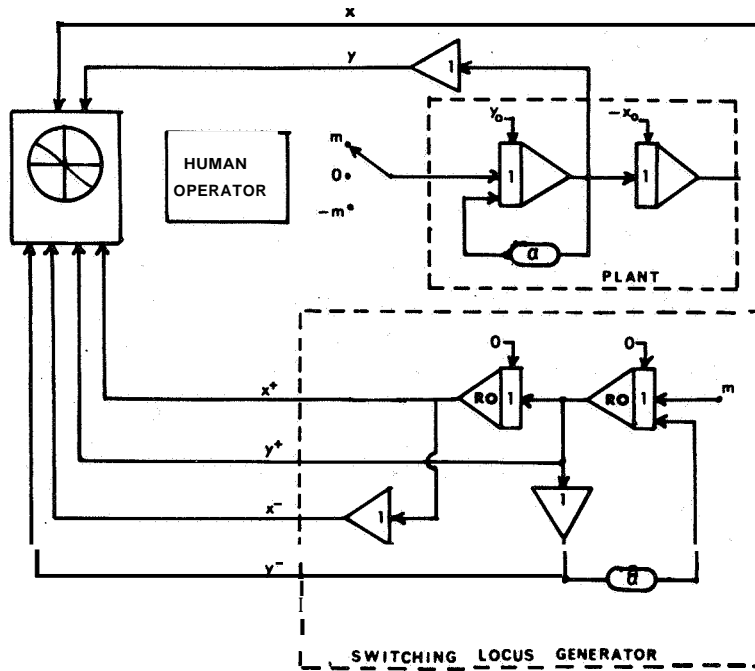


Figure 8.—Analog computer realization of manual time optimum control for $1/(s(s+0.16))$.

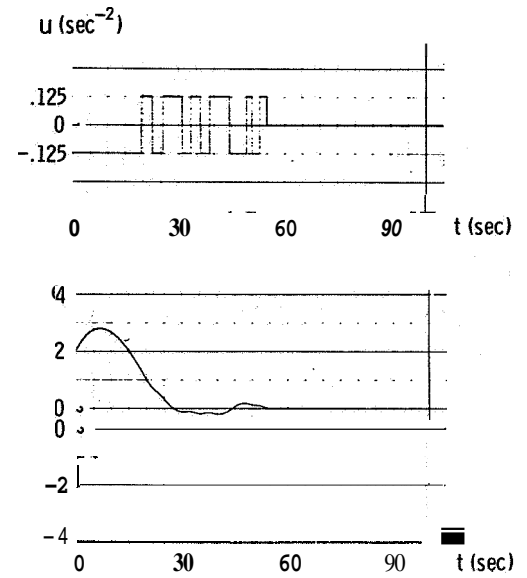


Figure 9.—A typical control force u and error e versus time obtained by experienced human operator stabilizing $1/(s(s+0.16))$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$), when state dot (e, \dot{e}) is displayed but no switching locus is displayed.

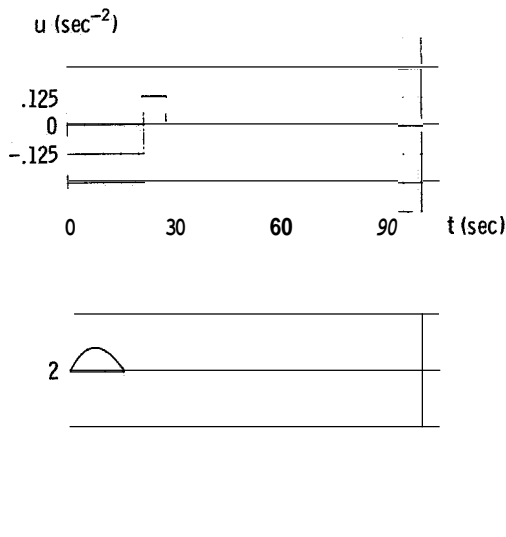


Figure 10.—A manual time optimum control result for $1/(s(s+0.16))$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$) when both state dot (e, \dot{e}) and exact optimum switching locus are displayed.

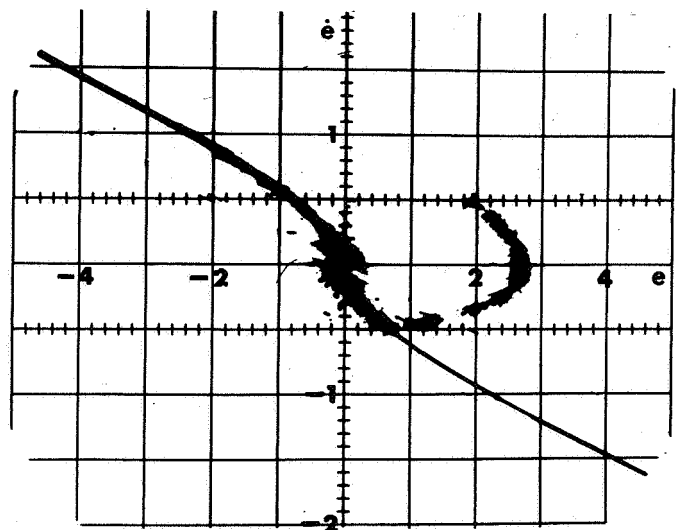


Figure 11.—Photograph of a time optimum trajectory and exact switching locus in e - \dot{e} display for $1/(s(s+0.16))$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$).

The analog computer realization of this approximate switching locus is shown in figure 12. Figure 13 shows the result of substituting the approximate switching locus for the exact switching locus in the same plant as in figure 10, with the same initial conditions. The corresponding trajectory and the approximate switching locus on the e- δ display are shown in figure 14. The state dot hit the approximate S_1^+ at 21 seconds, and the control reversal was made instantly. The state dot did not precisely follow S_1^+ after the reversal, since S_1^+ is approximate, but it traversed S_1^+ and then slightly deviated toward the negative error direction from the locus and hit S_1^- at 31 seconds. The second reversal was made, and the state dot followed S_1^- to the origin. The equilibrium state was reached at 33 seconds, 5 seconds later than with the exact switching locus.

EXAMPLES: $\frac{1}{s^3}$

The differential equation for the plant dynamics is:

$$\frac{d^3 e}{dt^3} = \pm m$$

The backward equation is then

$$\frac{d^3 e}{d\tau^3} = \mp m \text{ or } \begin{cases} \frac{dx}{d\tau} = -y \\ \frac{dy}{d\tau} = -z \\ \frac{dz}{d\tau} = \mp m \end{cases}$$

Integrating the above equation with respect to τ with null initial conditions yields:

$$x = \mp \frac{m}{6} \tau_1^3$$

$$y = \pm \frac{m}{2} \tau_1^2$$

$$z = \mp \tau_1$$

This set of parametric equations describes, in three-dimensional state space, a curve which was earlier defined as the switching subset, S_1 . By definition, $u = +m$ along S_1^+ and $u = -m$ along S_1^- . Therefore S_1^+ must correspond to the upper sign and S_1^- to the lower sign. As seen in figure 15, S_1^+ exists in the octant where $x \leq 0$, $y \geq 0$, and $z \leq 0$; S_1^- exists in the octant where $x \geq 0$, $y \leq 0$, and $z \geq 0$.

Next assume that the polarity of the control force is reversed at $\tau = \tau_1$. The backward equation after this reversal is

$$\frac{dx}{d\tau} = -y, \quad \frac{dy}{d\tau} = -z, \quad \frac{dz}{d\tau} = \pm m$$

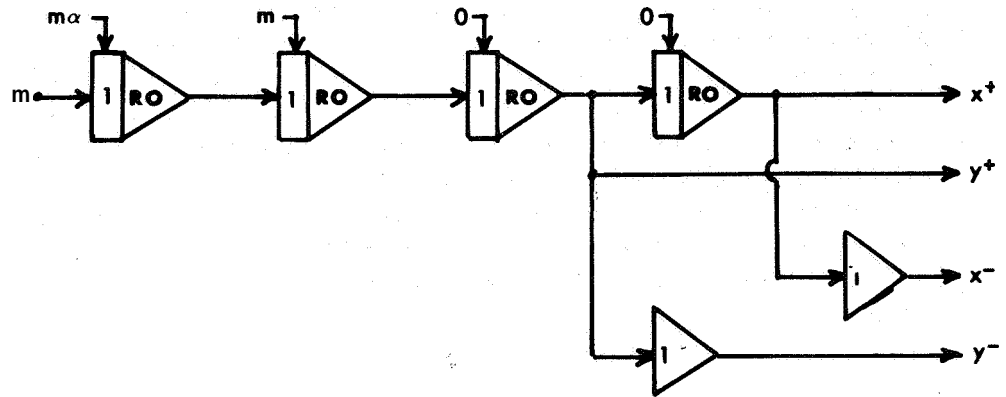


Figure 12.—Approximate switching locus generator for $1/(s(s+0.16))$.

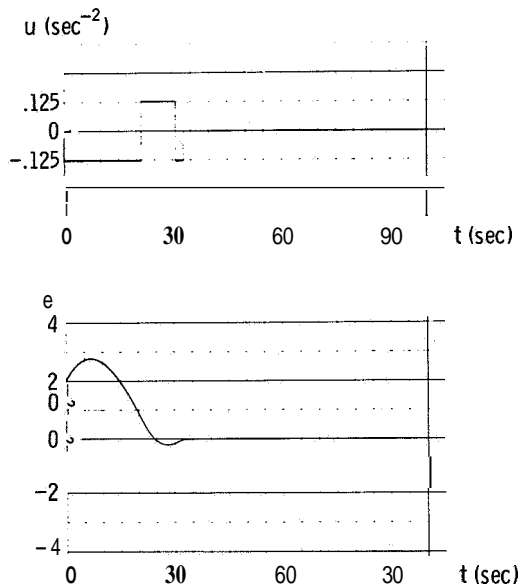


Figure 13.—A manual time optimum control result for $1/(s(s+0.16))$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$), when both state dot (e, \dot{e}) and approximate optimum switching locus are displayed.

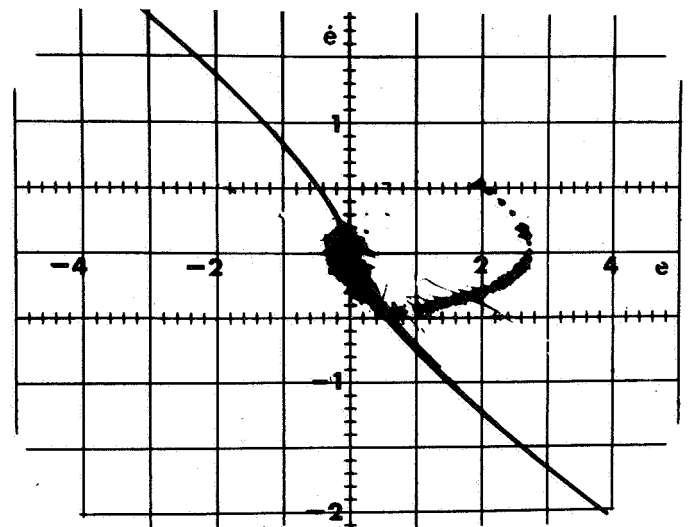


Figure 14.—Photograph of a quasi-time optimum trajectory and approximate switching locus in $e-\dot{e}$ display for $1/(s(s+0.16))$ with initial condition ($e=2.0$, $\dot{e}=0.5 \text{ sec}^{-1}$).

Integrating these equations with respect to τ from 0 to τ_2 , and taking the $x(\tau_1)$, $y(\tau_1)$, and $z(\tau_1)$ describing S_1 as initial conditions yields:

$$\begin{aligned}
 x &= \pm m \left(\frac{1}{6} \tau_2^2 - \frac{1}{2} \tau_2^2 \tau_1 - \frac{1}{2} \tau_2 \tau_1^2 - \frac{1}{6} \tau_1^3 \right) \\
 y &= \pm m \left(-\frac{1}{2} \tau_2^2 + \tau_1 \tau_2 + \frac{1}{2} \tau_1^2 \right) \\
 z &= \pm m (\tau_2 - \tau_1)
 \end{aligned}$$

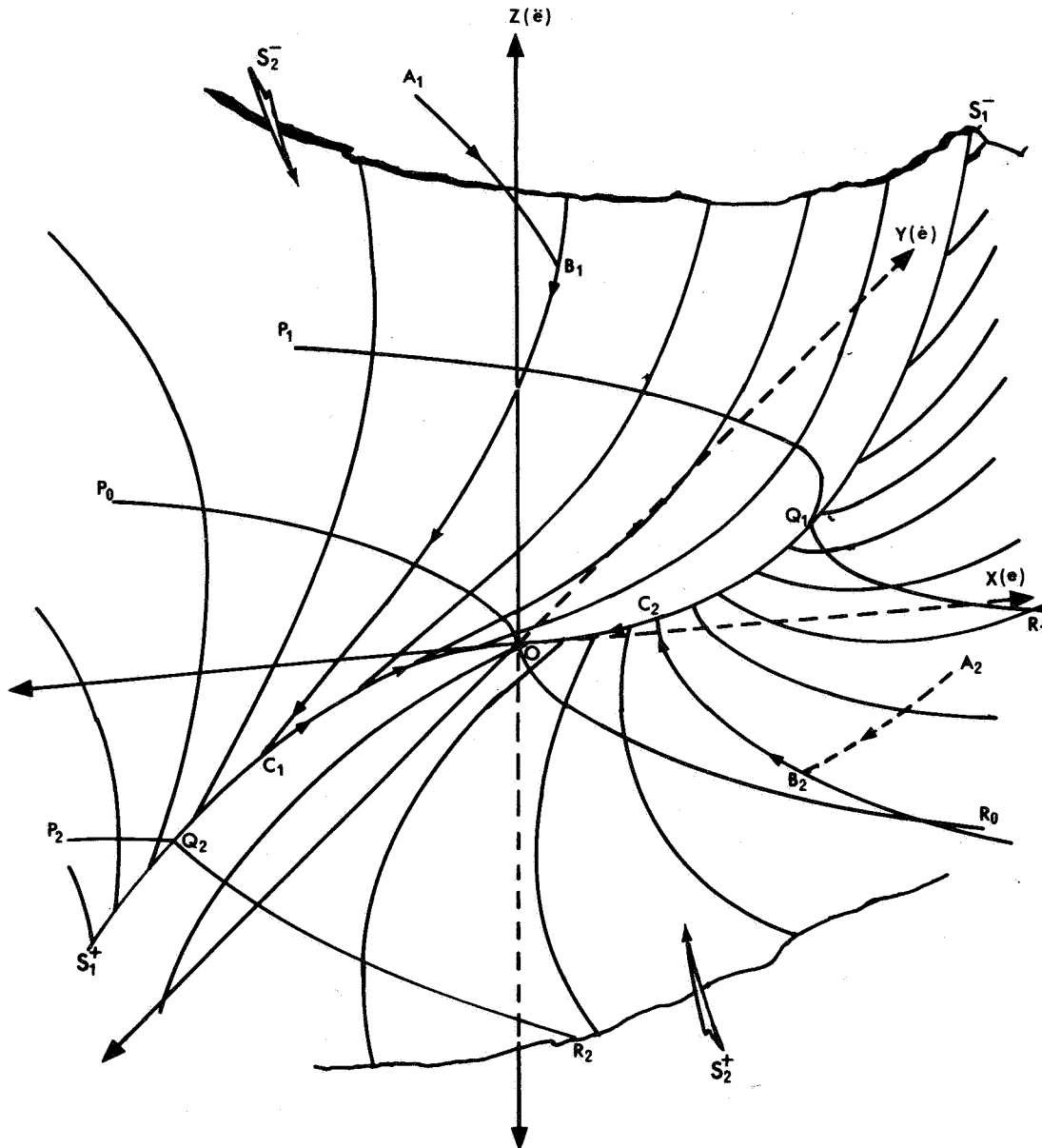


Figure 15.—Construction of optimum switching surface for $1/s^3$.

This set of parametric equations describes the switching surface S_2 , with S_2^+ corresponding to the upper sign, and S_2^- to the lower sign. S_2^+ and S_2^- are continuously connected, with S_1 as a boundary. This boundary gives the surface a corner. Note that the whole switching surface S_2 divides the entire state space into $+m$ control region S_3^+ and $-m$ control region S_3^- . In figure 15 the visible space corresponds to S_3^+ , while the space behind the switching surface corresponds to S_3^- . Any time-optimum trajectory of the state reaches the origin either via $S_3^+ \rightarrow S_2^- \rightarrow S_1^+ \rightarrow 0$, or via $S_3^- \rightarrow S_2^+ \rightarrow S_1^- \rightarrow 0$. The trajectory $A_1B_1C_10$ represents the former case and $A_2B_2C_20$ the latter. In order to derive the switching locus generating equation, τ_1 or τ_2 must be eliminated from the parametric equations representing the switching surface S_2 , and the remaining τ must be able to take values from 0. If

$$z = \pm m(\tau_2 - \tau_1), \text{ and } \tau_1 \geq 0$$

$$\tau_2 \geq 0$$

and $\tau_1 \geq 0$ and $\tau_2 \geq 0$, then for the upper sign (S_2^+):

if $z \geq 0$, τ_1 is the independent variable

if $z \leq 0$, τ_2 is the independent variable

For the lower sign (S_2^-):

if $z \geq 0$, τ_2 is the independent variable

if $z \leq 0$, τ_1 is the independent variable

It may be shown that such a choice is reasonable by examining the nature of a typical switching locus, $P_1Q_1R_1$ for $z > 0$, $P_2Q_2R_2$ for $z < 0$, and $P_0Q_0R_0$ for $z = 0$. The generating equations are now:

For the upper sign (S^+):

if $z \geq 0$

$$x = m\left(-\tau_1^3 - z\tau_1^2 + \frac{z^3}{6}\right)$$

$$y = m\left(\tau_1^2 - \frac{z^2}{2}\right)$$

if $z \leq 0$

$$x = m\left(-\tau_1^3 + 2z\tau_1^2 - z^2\tau_1 + \frac{z^3}{6}\right)$$

$$y = m\left(\tau_1^2 - 2z\tau_1 + \frac{z^2}{2}\right)$$

For the lower sign (S^-):

if $z \geq 0$

$$x = m\left(\tau_1^3 + 2z\tau_1^2 + z^2\tau_1 + \frac{z^3}{6}\right)$$

$$y = m\left(-\tau_1^2 - 2z\tau_1 - \frac{z^2}{2}\right)$$

if $z \leq 0$

$$x = m\left(\tau_1^3 - \frac{z^3}{6}\right)$$

$$y = m\left(-\tau_1^2 + \frac{z^2}{2}\right)$$

The corresponding differential equations and associated initial conditions are obtained as shown in table 1. Figure 16 shows the computer realization of the exact switching locus for a manual-control experiment with this $1/s^3$ plant. The control force chosen was $m=0.1$ second $^{-3}$. Two initial conditions were used:

$$A: (0.025e=1.0, 0.25\dot{e}=0.5 \text{ sec}^{-1}, \ddot{e}=0.5 \text{ sec}^{-2})$$

$$B: (0.025e=0.5, 0.25\dot{e}=-1.0 \text{ sec}^{-1}, \ddot{e}=-0.5 \text{ sec}^{-2})$$

TABLE 1.—OPTIMUM SWITCHING LOCUS GENERATING EQUATION
WITH CORRESPONDING INITIAL CONDITIONS

| Upper sign | Lower sign |
|--|---|
| $\frac{d^3x}{dt^3} = -6m$ | $\frac{d^3x}{dt^3} = -6m$ |
| $\frac{d^2y}{dt^2} = 2m$ | $\frac{d^2y}{dt^2} = -2m$ |
| <p>If $z \geq 0$:</p> $x(0) = x^3/6m^2$ $\dot{x}(0) = 0$ $\ddot{x}(0) = -2z$ $y(0) = -z^2/2m$ $\dot{y}(0) = 0$ <p>If $z \leq 0$:</p> $x(0) = z^3/6m^2$ $\dot{x}(0) = -z^2/m$ $\ddot{x}(0) = 4z$ $y(0) = z^2/2m$ $\dot{y}(0) = -2z$ | <p>If $z \geq 0$:</p> $x(0) = z^3/6m^2$ $\dot{x}(0) = z^2/m$ $\ddot{x}(0) = 4z$ $y(0) = -z^2/2m$ $\dot{y}(0) = -2z$ <p>If $z \leq 0$:</p> $x(0) = z^3/6m^2$ $\dot{x}(0) = 0$ $\ddot{x}(0) = -2z$ $y(0) = z^2/2m$ $\dot{y}(0) = 0$ |

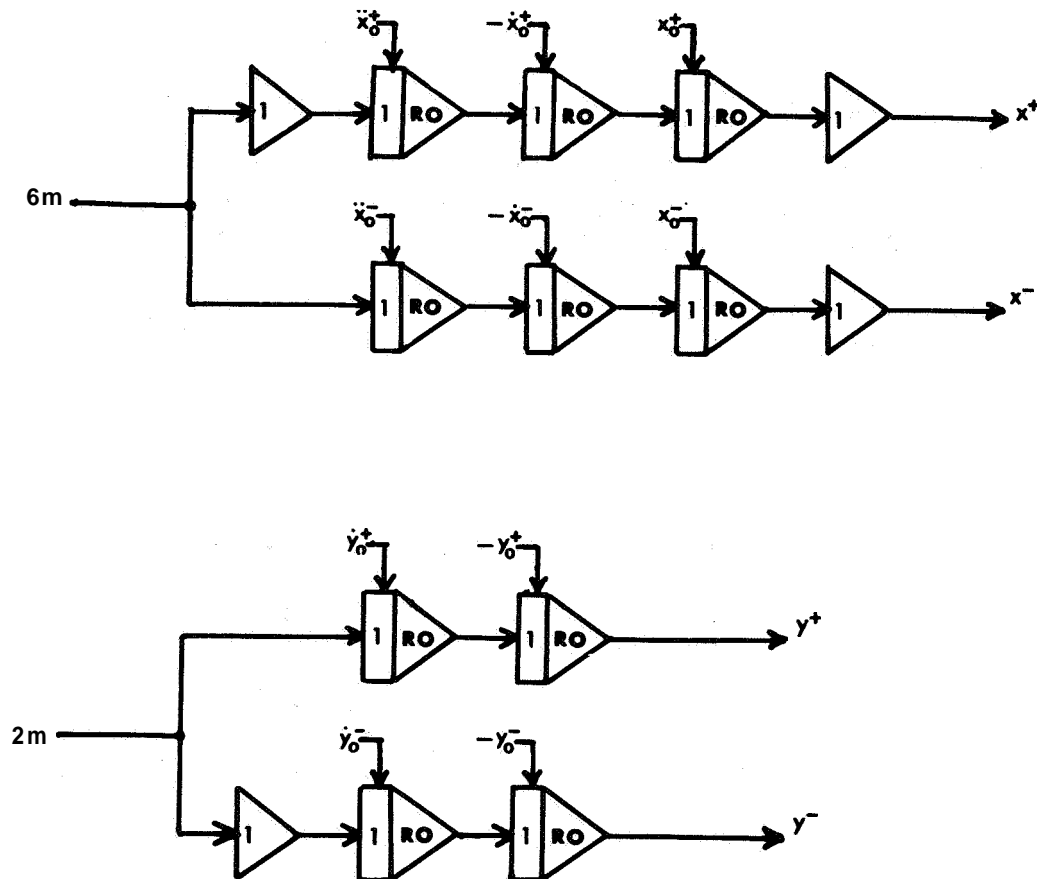
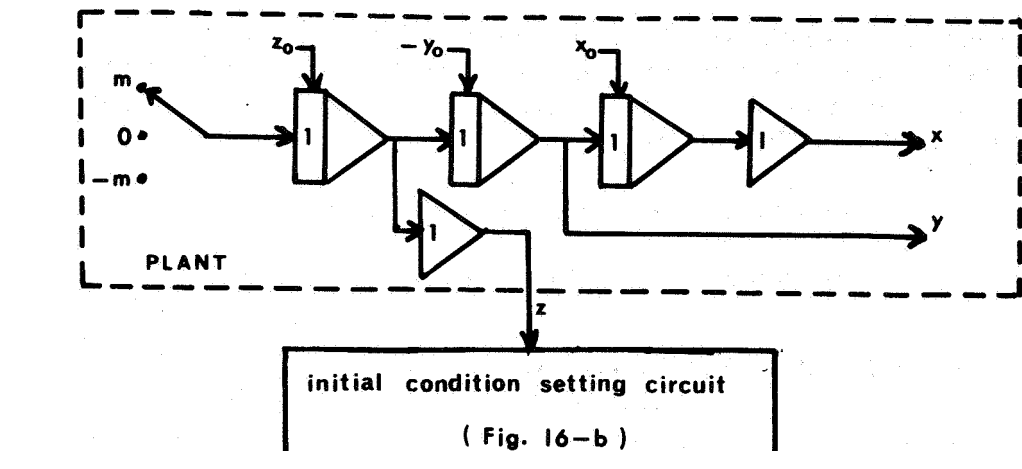


Figure 16a.—Analog computer realization of exact optimum switching locus generator for $1/s^3$.

Figure 17 and figure 18 show typical results obtained by an experienced operator with no switching locus for initial conditions A and B, respectively. The chatter mode is clearly

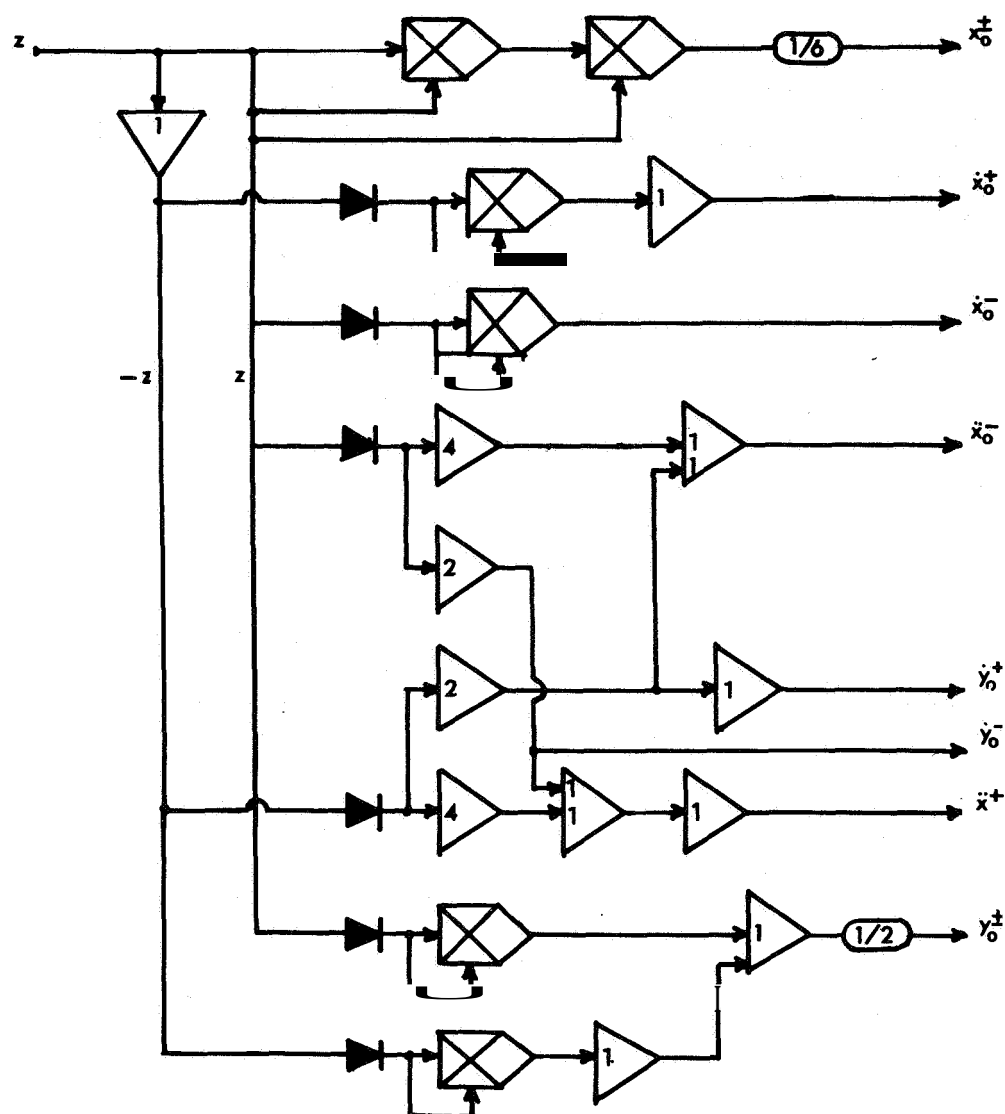


Figure 16b.—Initial condition setting circuit.

observed to approach the equilibrium state in both figures. Figures 19 and 20 show "time-optimum manual control" results for A and B, respectively, when the exact switching locus was displayed. The time necessary to reach the equilibrium state **for** each case was reduced to less than half the time required when no switching locus was displayed. The number of reversal times is seen to be three **for** A and four for B, both of which are more than the minimum of two switchings that the theory would predict. The reasons for **this** are human timing errors and instrumentation inaccuracy of the switching locus generator mainly due to the diode multipliers used in the analog computer. Figures 21 and 22 illustrate the sequential events on the $e-\dot{e}$ display for A and B, respectively. Tables 2 and 3 interpret the photographs in figures 21 and 22, respectively.

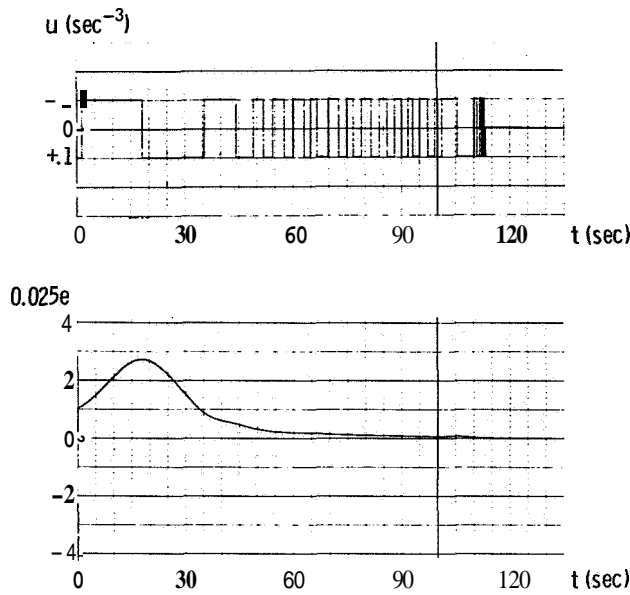


Figure 17.—A typical control force u and error e versus time obtained by experienced human operator stabilizing $1/s^3$ with initial condition ($0.025e=1.0$, $0.25\dot{e}=0.5 \text{ sec}^{-1}$, $\ddot{e}=0.5 \text{ sec}^{-2}$), when state dot is displayed but no switching locus is displayed.

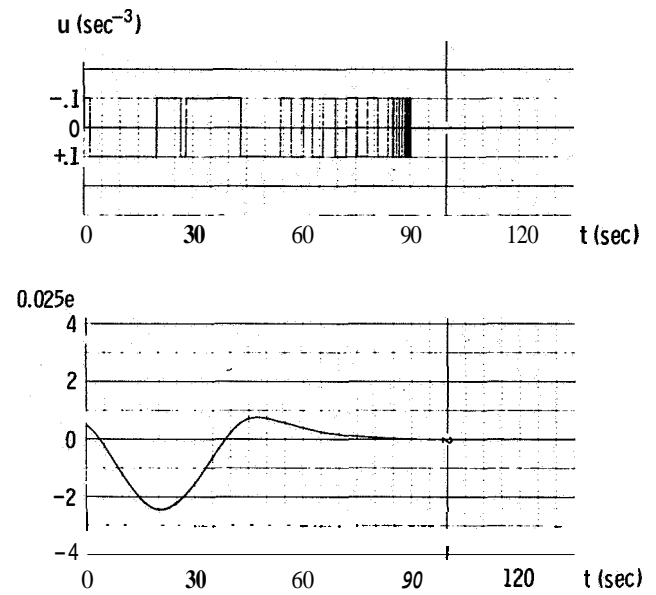


Figure 18.—A typical control force u and error e versus time obtained by experienced human operator stabilizing $1/s^3$ with initial condition ($0.025e=0.5$, $0.25\dot{e}=-1.0 \text{ sec}^{-1}$, $\ddot{e}=-0.5 \text{ sec}^{-2}$), when state dot is displayed but no switching locus is displayed.

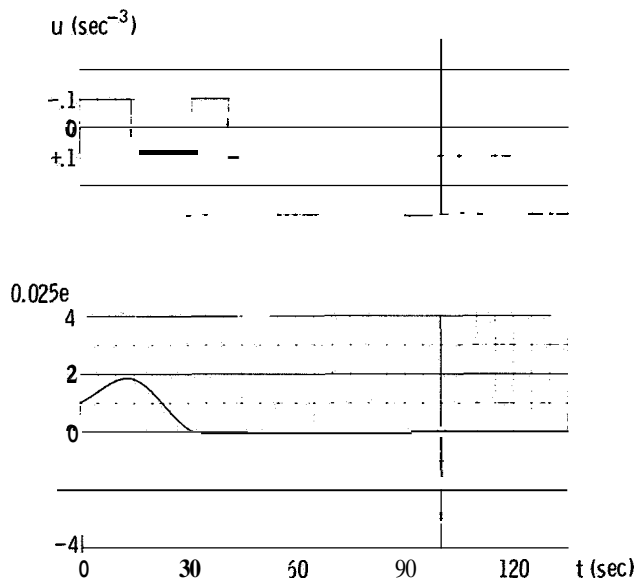


Figure 19.—A manual time optimum control result for $1/s^3$ with initial condition ($0.025e=1.0$, $0.25\dot{e}=0.5 \text{ sec}^{-1}$, $\ddot{e}=0.5 \text{ sec}^{-2}$) when both state dot (e, \dot{e}) and exact optimum switching locus are displayed.

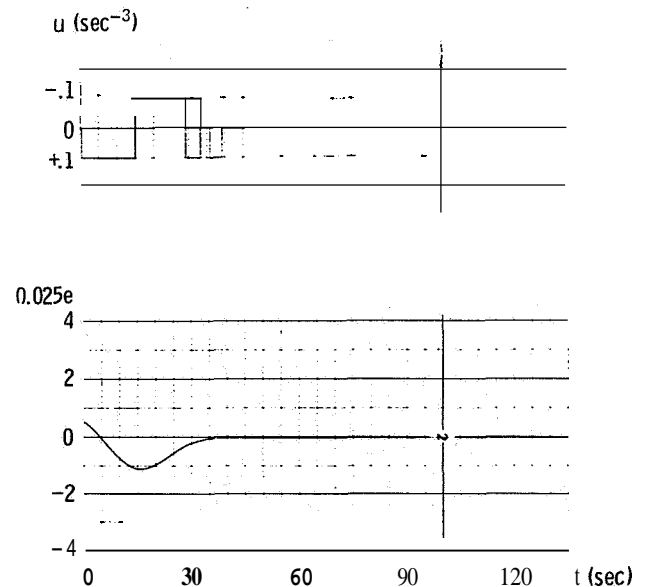


Figure 20.—A manual time optimum control result for $1/s^3$ with initial condition ($0.025e=0.5$, $0.25\dot{e}=-1.0 \text{ sec}^{-1}$, $\ddot{e}=-0.5 \text{ sec}^{-2}$), when both state dot (e, \dot{e}) and exact optimum switching locus are displayed.

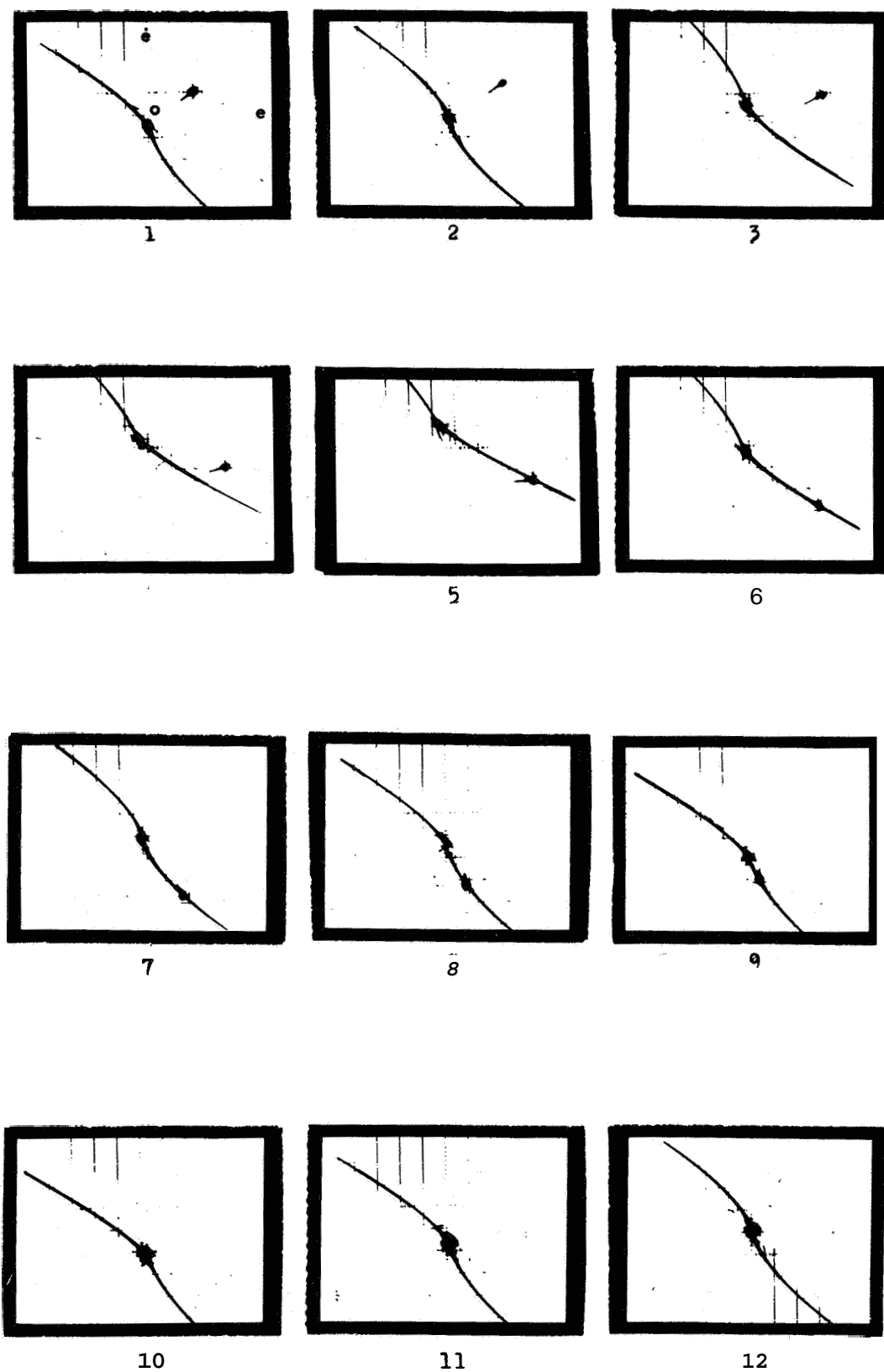


Figure 21.—Sequential photographs of $e-\dot{e}$ display with exact switching locus for $1/s^3$ with initial condition ($0.025e=1.0$, $0.25\dot{e}=0.5$ sec, $\ddot{e}=0.5$ sec).

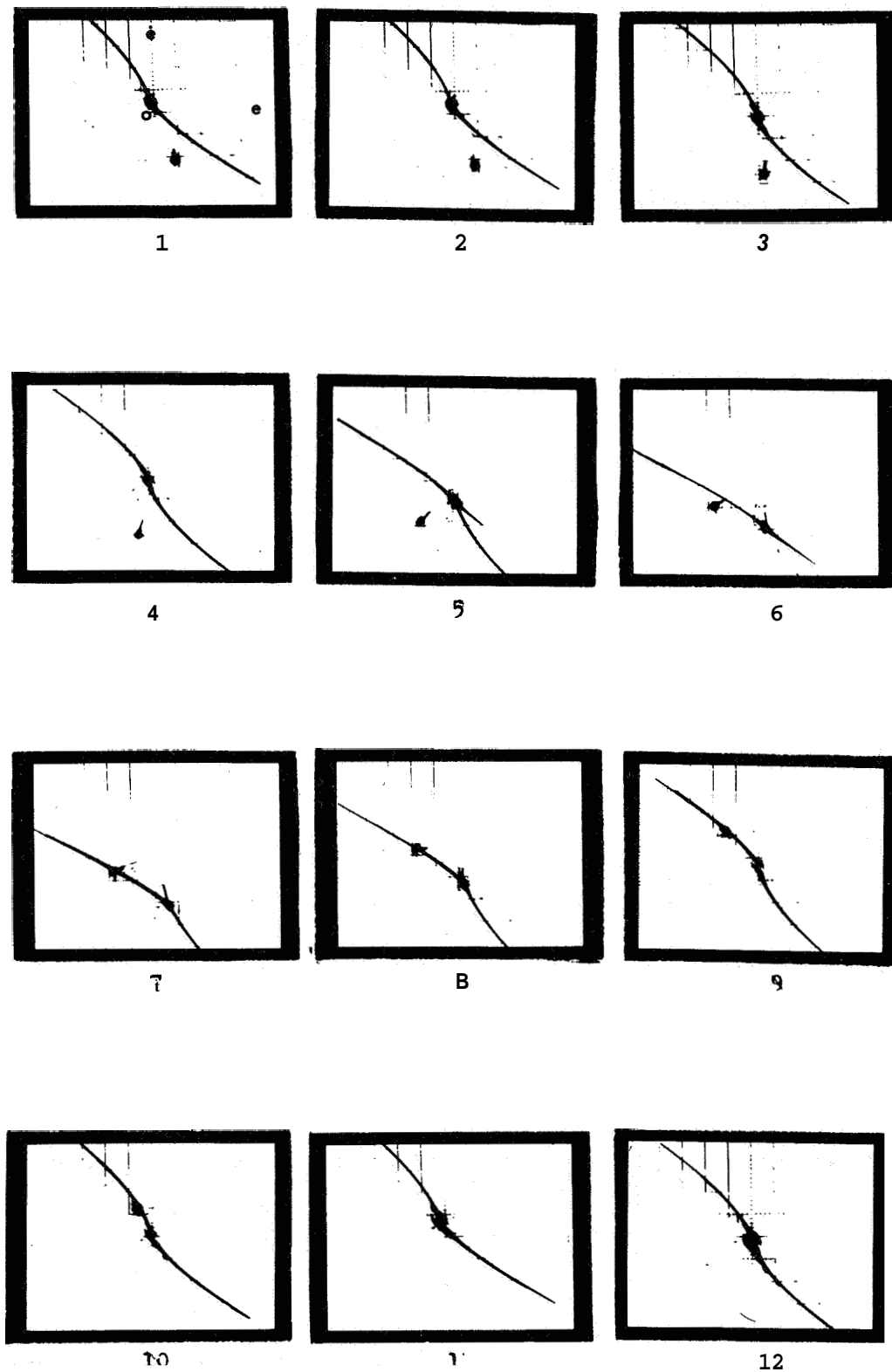


Figure 22.—Sequence photographs of $e-\dot{e}$ display with exact switching locus for $1/s^3$ with initial condition ($0.025e=0.5$, $0.25\dot{e}=-1.0 \text{ sec}^{-1}$, $\ddot{e}=-0.5 \text{ sec}^{-2}$).

TABLE 2.—INTERPRETATION OF SEQUENTIAL PHOTOGRAPHS
IN FIGURE 21

| Photograph | Event | Control polarity | Region of state dot |
|------------|----------------------------------|------------------|---------------------|
| 1 | Initial state | 0 | S_3^- |
| 2 | | - | S_3^- |
| 3 | | - | S_3^- |
| 4 | | - | S_3^- |
| 5 | Encounter with switching surface | Switching | S_2^+ |
| 6 | | + | S_2^+ |
| 7 | | + | S_2^+ |
| 8 | | + | S_2^+ |
| 9 | | + | S_2^+ |
| 10 | Encounter with corner curve | Switching | S_1^+ |
| 11 | | - | S_1^- |
| 12 | Equilibrium state | 0 | Origin |

TABLE 3.—INTERPRETATION OF SEQUENTIAL PHOTOGRAPHS
IN FIGURE 22

| Photograph | Event | Control polarity | Region of state dot |
|------------|----------------------------------|------------------|---------------------|
| 1 | Initial state | 0 | S_3^+ |
| 2 | | + | S_3^+ |
| 3 | | + | S_3^+ |
| 4 | | + | S_3^+ |
| 5 | | + | S_3^+ |
| 6 | | + | S_3^+ |
| 7 | Encounter with switching surface | - | S_2^- |
| 8 | | - | S_2^- |
| 9 | | - | S_2^- |
| 10 | | - | S_2^- |
| 11 | Encounter with corner curve | + | S_1^+ |
| 12 | Equilibrium state | 0 | Origin |

CONCLUSION

A visual display permitting manual, time-optimal bang-bang control is possible with the familiar $e-\dot{e}$ plane, on which the instantaneous $e-\dot{e}$ and the "switching locus" are simultaneously displayed. Error and error-rate information are constantly available to the human operator in this closed-loop control system.

For any second-order plant with real poles and for one third-order plant, $1/s^3$, the implementation of such a system is not difficult. For an arbitrary higher order plant, however, the analytical description of the switching surface is usually difficult to obtain and too complicated to instrument even if obtained. Some approximation for the switching subspace of such plants may thus be required for its realization. Either a Maclaurin expansion of exponential terms involved in the analytical expression of the switching subspace, or the least squared error fitting of an assumed expression as against the exact one may be used for the approximation. The display works reasonably well with either approximation in a certain limited region of the state space.

REFERENCES

1. Platzer, Harold L.: The Phase Plane as a Tool for the Study of Human Behavior in Tracking Problems. WADC Tech. Rep. 55-444, Nov. 1955.
2. Platzer, Harold L.: The Non-Linear Approach to Human Tracking. Interim tech. rep. no. 1-2490-1, Laboratories for Research and Development, The Franklin Institute, Dec. 21, 1955.
3. Li, Yao Tzu; Young, L. R.; and Meiry, J. L.: Adaptive Function of Man in Vehicle Control Systems. IFAC Symposium, Sept. 1965.
4. Chang, Sheldon S. L.: Synthesis of Optimum Control Systems. McGraw-Hill Book Co., 1961.
5. Athans, Michael; and Falb, Peter L.: Optimal Control. McGraw-Hill Book Co., 1966.
6. Ghandhuri, A. K.; and Choudhury, A. K.: On the Optimum Switching Function of a Certain Class of Third-Order Contactor Servomechanism I. J. of Electron. and Control, vol. 16, no. 4, Apr. 1964.
7. Ghandhuri, A. K.; and Choudhury, A. K.: On the Optimum Switching Function of a Certain Class of Third-Order Contactor Servomechanism II. J. of Electron. and Control, vol. 17, no. 4, Oct. 1964.

28. Adaptive Finite-State Models of the Human Operator"

E. S. Angel
University of Southern California

N68-15929

While most of the mathematical models of human operators are based on the operator acting in a continuous manner upon continuous data, this model is based upon the human operator seeing only quantized input data and possessing a small number of internal states. The basic model is shown here, and a scheme by which the threshold levels might be adjusted to make the basic model adaptive is presented. Some preliminary results and suggestions for further research are also presented.

Most of the present mathematical models of human operators are based on the operator acting in a continuous manner upon continuous input data. The model presented here is based on the assumption that the human operator possesses only a small number of internal states and changes states on the basis of quantized observations of error and error rate.

The basic model is that previously proposed by Bekey and Angel (ref. 1). The model uses the concept of "force programs" (refs. 2 and 3) (prestored error correction patterns) to give specific responses of the model based on its inputs and internal state. Continuous outputs are obtained by the use of "hybrid actuators" (ref. 4).

The original model, while only part of a feasibility study and not intended to match closely real human operators, possessed some of the important characteristics of human operators. Specifically, the responses were of finite duration and could not be interrupted until an action in progress had run to completion. Furthermore, the model was able to track precisely nonaccelerating inputs. However, the model was not adaptive; it could not improve its performance over longer time of observing the same input curve.

BASIC MODEL

The basic model was designed to simulate a human operator in a compensatory tracking task with a pure inertia plant. The system is as shown in figure 1. The operator sees only error and error rate. He quantizes these quantities, and on the basis of these threshold levels and his present state he generates a force program.

In this simple model the operator can do one of three things if he is in a state where he can make a decision, that is, not already in the middle of a force program: he can do nothing, he can attempt to change his position, or he can attempt to change his velocity. In this model

"This research was sponsored in part by the National Aeronautics and Space Administration under Grant NGR-05-018-022 and in part by the U. S. Air Force Office of Scientific Research under Grant AF-AFOSR-1018-67.

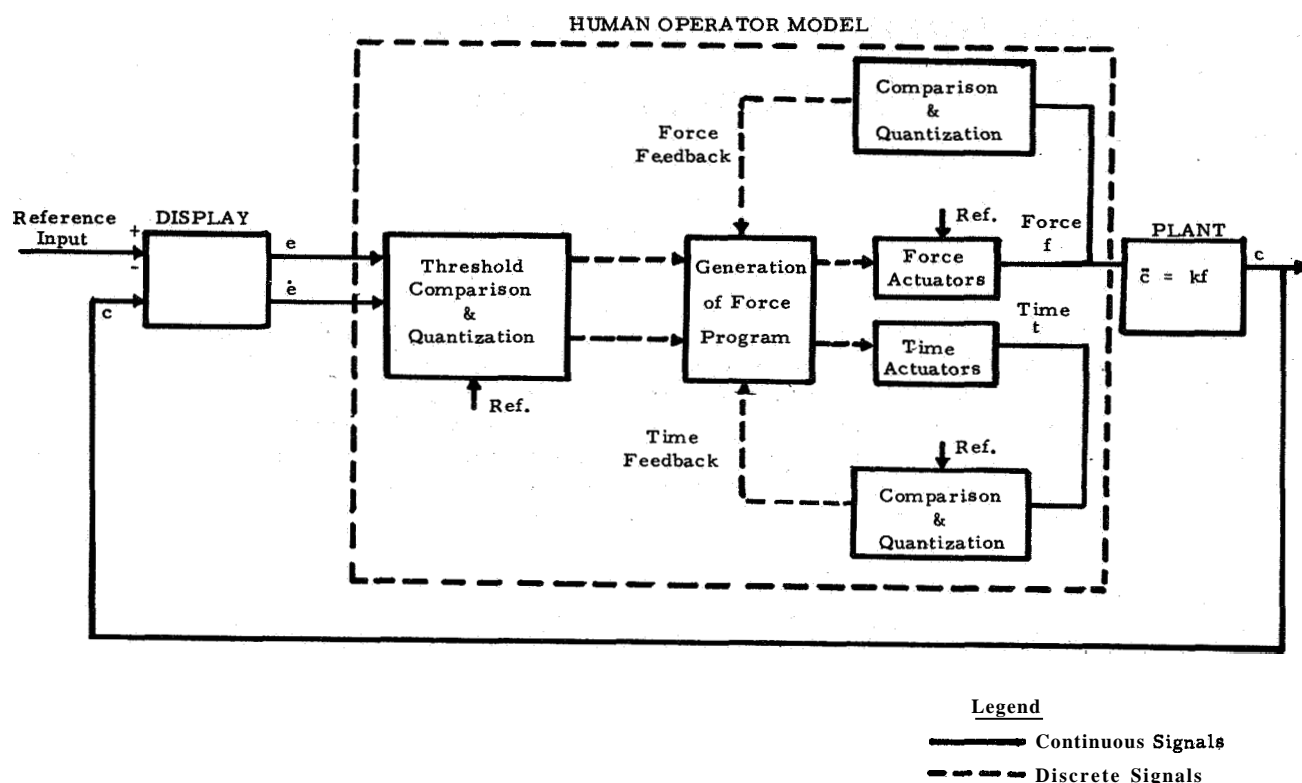


Figure 1.—Block diagram of basic model.

no memory was used so the decision was made on the basis of combinations as shown in table 1. Thus for the two error thresholds and one rate threshold model, decisions are made on the basis of this table. In this model, only two position and one velocity correction (and their negatives) were allowed. The structure of this process is shown in figure 2. The time actuators insure that the force program has the proper duration. The model can be simulated with only a very small number of logical gates, actuators, threshold gates, and flip flops.

TABLE 1.—COMBINATIONS USED FOR DECISION-MAKING

| $-\epsilon_2$ | $-\epsilon_1$ | 0 | 0 | ϵ_1 | ϵ_2 | |
|---------------|---------------|----|----|--------------|--------------|-------------|
| +p | +p | -v | -v | -v | -p | \dot{e} |
| +p | +p | 0 | 0 | -p | -p | 0 |
| +p | +p | 0 | 0 | -p | -p | 0 |
| +p | +p | +v | +v | -p | -p | $-\epsilon$ |

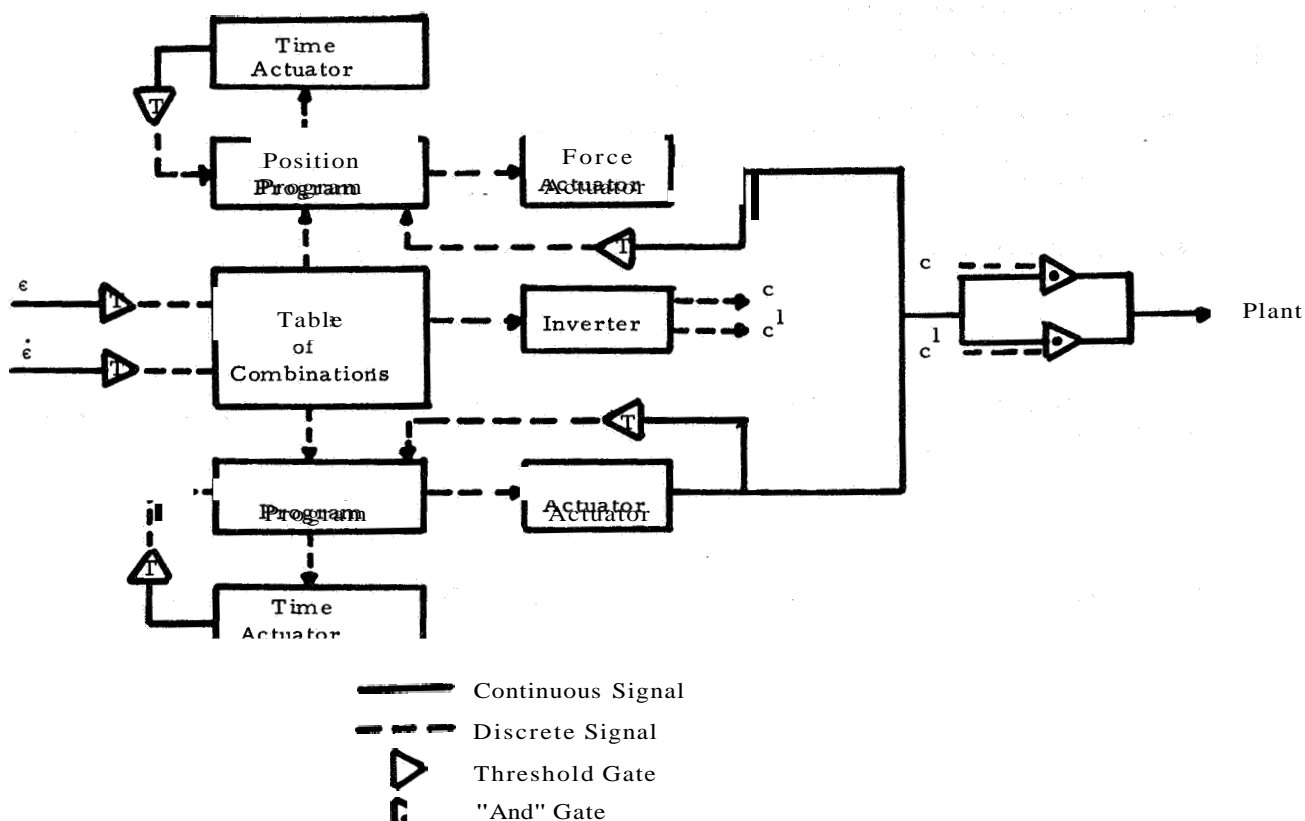


Figure 2.—Block diagram of asynchronous network.

REQUIREMENTS OF AN ADAPTIVE MODEL

An attempt was made to alter the basic model so that it would be able to improve its performance over time. We want to achieve this not by adding many more threshold levels but by having the ability to adjust the few levels used in the basic model.

Some of the necessary features are:

- (1) Ability to reduce error and error rate to zero for a "simple" input curve
- (2) Ability to adjust to a change in input curve
- (3) Ability to get "close" to zero errors after three or four corrections

ADJUSTMENT OF THE THRESHOLDS

We desire a device which, on the basis of error and error rate thresholds, will decide by use of a finite state machine to make one of a finite number of position or velocity corrections and/or to adjust its threshold levels in such a way as to reduce the error and error rate to zero. The thresholds are to be placed so that if the error e and the error rate \dot{e} remain constant for the duration of the correction, the error (error rate) if it was corrected will be smaller than the smallest error (error rate) threshold.

Let

n number of positive velocity thresholds

m number of positive position thresholds

e_i level of i th position threshold

\dot{e}_j level of j th position threshold

f_i correction for $e_{i+1} > |e| > e_i$

\dot{f}_j correction for $\dot{e}_{j+1} > |\dot{e}| > \dot{e}_j$

Consider first the velocity thresholds. We have $2n+1$ levels to consider:

$$\dot{e}_n > \dot{e}_{n-1} > \dots > \dot{e}_1 > 0 > -\dot{e}_1 > \dots > -\dot{e}_n$$

We wish to map all points of the velocity error \dot{e} into the region $(\dot{e}_1, -\dot{e}_1)$. Thus we seek the set $\{f_i\}$ such that:

$$\dot{e}_1 \geq |\dot{e}| - f_i \geq -\dot{e}_1, \text{ for } \dot{e}_{i+1} \geq |\dot{e}| \geq \dot{e}_i$$

This is illustrated in figure 3 for a number of possible corrections. In particular, consider the point \dot{e}_{i+1} . The corrections f_i and \dot{f}_{i+1} must be chosen so that

$$\dot{e}_1 \geq \dot{e}_{i+1} - f_{i+1} \geq -\dot{e}_1 \text{ and } \dot{e}_1 \geq \dot{e}_{i+1} - \dot{f}_{i+1} \geq -\dot{e}_1$$

since \dot{e}_{i+1} is a boundary point.

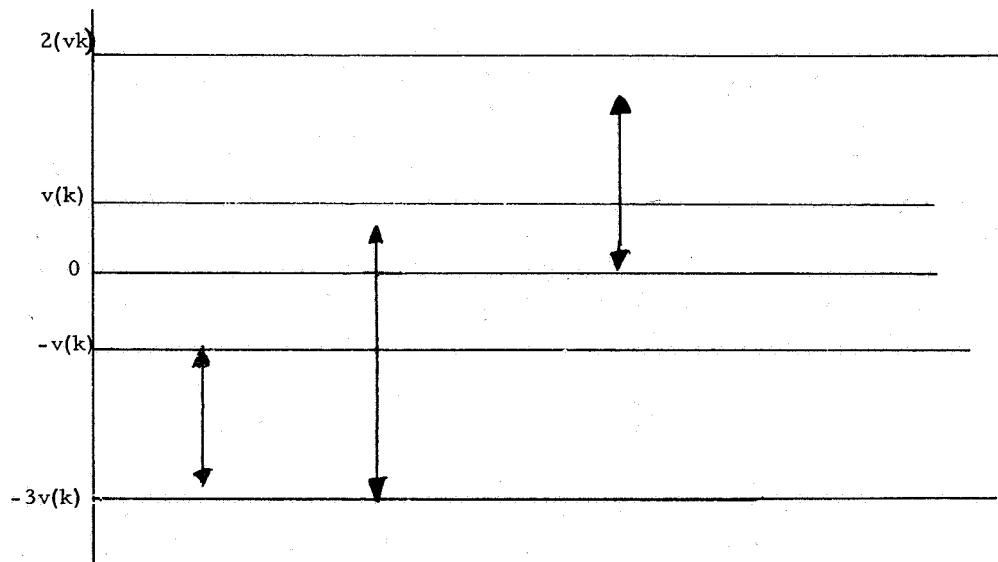


Figure 3.—Velocity threshold levels.

If every point in $(\dot{e}_1, -\dot{e}_1)$ is to be used since we want the most efficient scheme, the only solutions to the above equations are:

$$\dot{e}_{i+1} - f_i = \dot{e}_1$$

$$\dot{e}_{i+1} - f_{i+1} = -\dot{e}_1$$

$$\dot{e}_i - f_i = \dot{e}_1$$

$$f_i = \dot{e}_1 + \dot{e}_{i+1} - \dot{e}_1$$

$$\dot{e}_{i+1} = \dot{e}_1 + 2\dot{e}_1$$

The solution to this set of equations is

$$\left. \begin{aligned} \dot{e}_i &= (2i-1)\dot{e}_1 \\ f_i &= 2(i-1)\dot{e}_1 \end{aligned} \right\} (i=2, \dots, n)$$

Consider now the position correlations. The problem is the same as above except for the fact that the error in velocity may cause an additional position error of $\pm \dot{e}_1 \Delta t$, where Δt is the duration of a position correction (fig. 4). Thus we obtain for this case

$$\left. \begin{aligned} e_j &= (2j-1)(e_1 - \dot{e}_1 \Delta t) \\ f_j &= 2(j-1)(e_1 - \dot{e}_1 \Delta t) \end{aligned} \right\} (j=2, \dots, n)$$

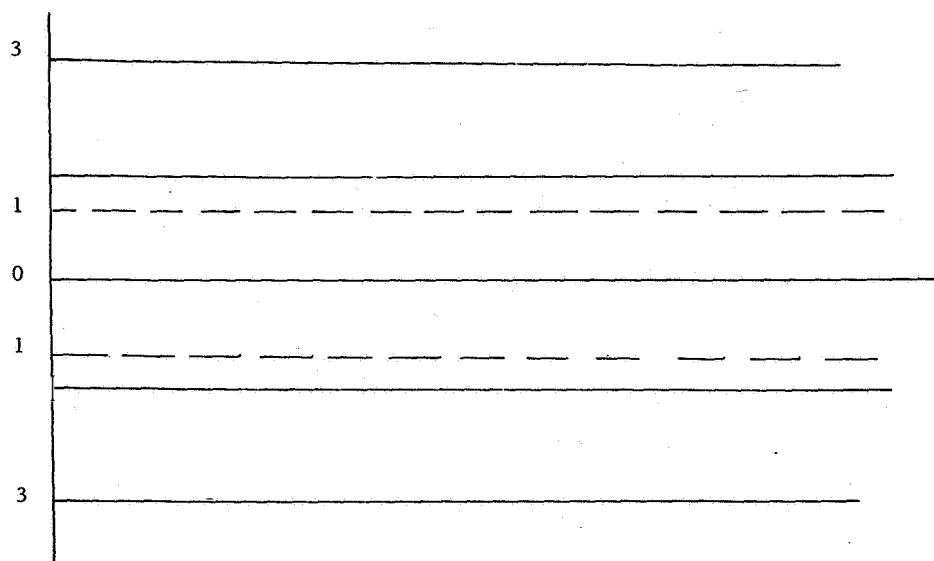


Figure 4.—Position threshold levels.

Thus with this correction scheme all errors are mapped into the region determined by the smallest thresholds. At this point the threshold levels are changed. Since the error is located in the innermost region, we just break up this region as before.

Let

k number of threshold adjustments

$e_v(k)$ \dot{e}_1 after k th adjustment

$e_x(k)$ e_1 after k th adjustment

Thus, we obtain the following set of equations if we allow n velocity thresholds and m position thresholds

$$e_v(k) = \frac{e_v(k-1)}{n}$$

$$e_x(k) = \frac{e_x(k-1)}{m} + e_v(k)\Delta t$$

Solving these two equations we obtain

$$e_v(k) = \frac{e_v(0)}{n^k}$$

$$e_x(k) = \frac{e_x(0)}{m^k} + \frac{e_v(0)\Delta t}{1 - \frac{n}{m}} \left(\frac{1}{n^k} - \frac{1}{m^k} \right) \quad (m \neq n)$$

$$e_x(k) = \frac{e_x(0) + k e_v(0)\Delta t}{n^k} \quad (m = n)$$

In order that the model operates correctly, we must have $e_1 < e_2$. This is not always guaranteed because of the $\dot{e}_1\Delta t$ term. For the case that $m = n$, the following condition will insure that this is true for all k :

$$e_x(0) > 3/2 e_v(0)\Delta t$$

It should be pointed out that the index k only denotes the k th adjustment but does not indicate at what time this adjustment occurs. This makes the adjustment scheme different from a sampled data or synchronous adjustment scheme.

THE ADAPTIVE MODEL

The adaptive model is obtained by using the basic model with the adjustment procedure of figure 5. Let $T(k)$ be the set of all threshold levels after the k th adjustment. The model then works as follows:

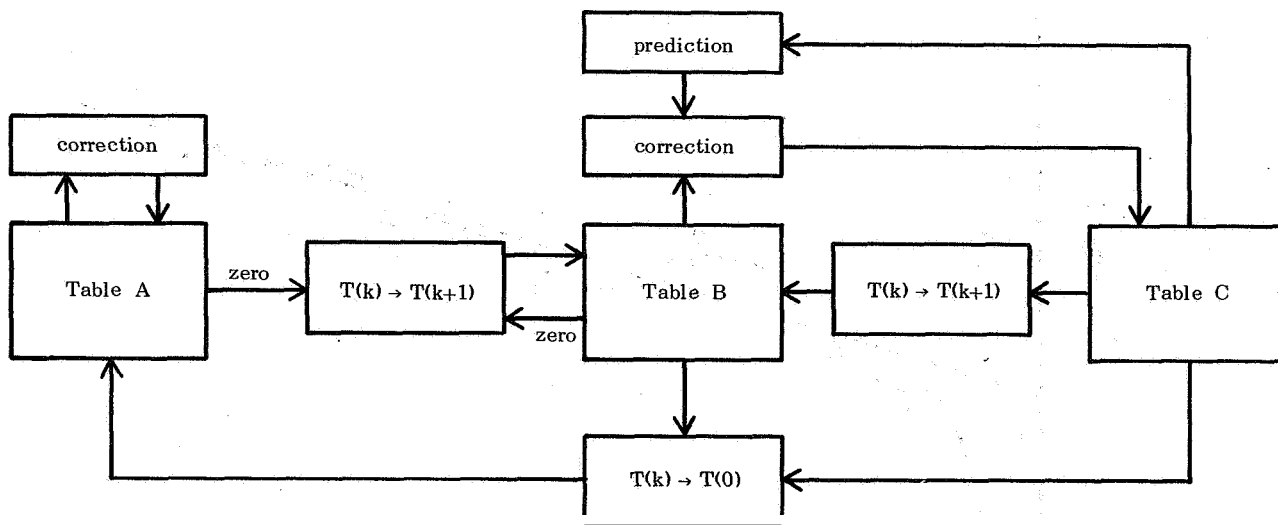


Figure 5.—Level adjustment scheme.

(1) On the basis of the original threshold levels, $T(0)$, and an initial table of combinations, table A, the model tries to reduce errors to zero. This corresponds to some initial rough corrections. When the errors reach the zero region, the rough correction stage is concluded.

(2) The threshold levels are adjusted to their next level, $T(1)$, and the same table of combinations or a new one, table B, is used. Here one of three things can happen. First, the errors can be so small that we are already in the zero region of table B. In this case we adjust the thresholds again, $T(k) \rightarrow T(k+1)$, and go back to table B. On the other hand, the errors might be so large that on the basis of some determined criteria, we assume that the input has changed; In this case, we change the thresholds to their original levels $T(0)$ and start again with table A. Finally, if neither of these things have happened, we make a correction on the basis of table B.

(3) After the correction is made, we check to see the result of it by table C. Remembering that the threshold levels and corrections were chosen so that for simple inputs the right correction would put the errors into the zero region, we have three possibilities to check for in table C. First, the correction could have done what it was intended to do; in this case the thresholds are again adjusted, $T(k) \rightarrow T(k+1)$, and we return to table B. Second, the input could have changed during the correction in which case we return to table A and the original threshold levels. Third, the correction has not reduced the errors into the zero region which means that the input is accelerating. In this case we assume that it will continue to accelerate and we use a correction which predicts where the input will be at the end of the correction time. Here, the amount of the correction is changed by a constant which is dependent upon the duration of the correction.

RESULTS

Some typical results of the adaptive model are shown in figures 6 and 7. The points at which corrections are begun are denoted by arrows.

For the case of simple inputs, for example, ramps and steps, the rate of convergence is almost entirely dependent upon the number of thresholds allowed. The specific table of combinations makes almost no difference at all.

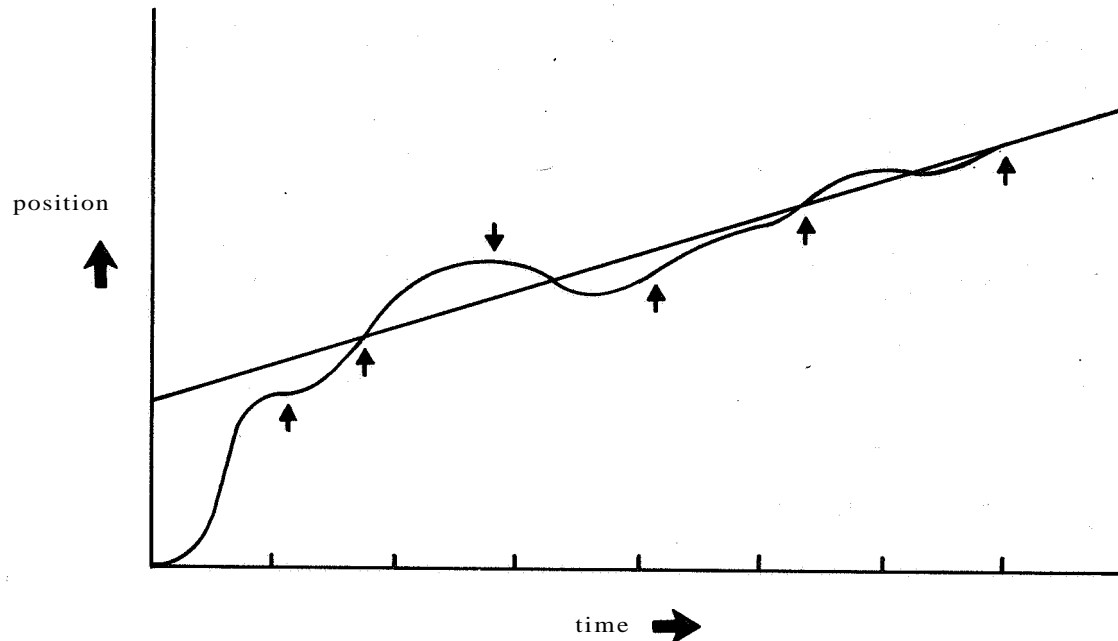


Figure 6.—Response to ramp.

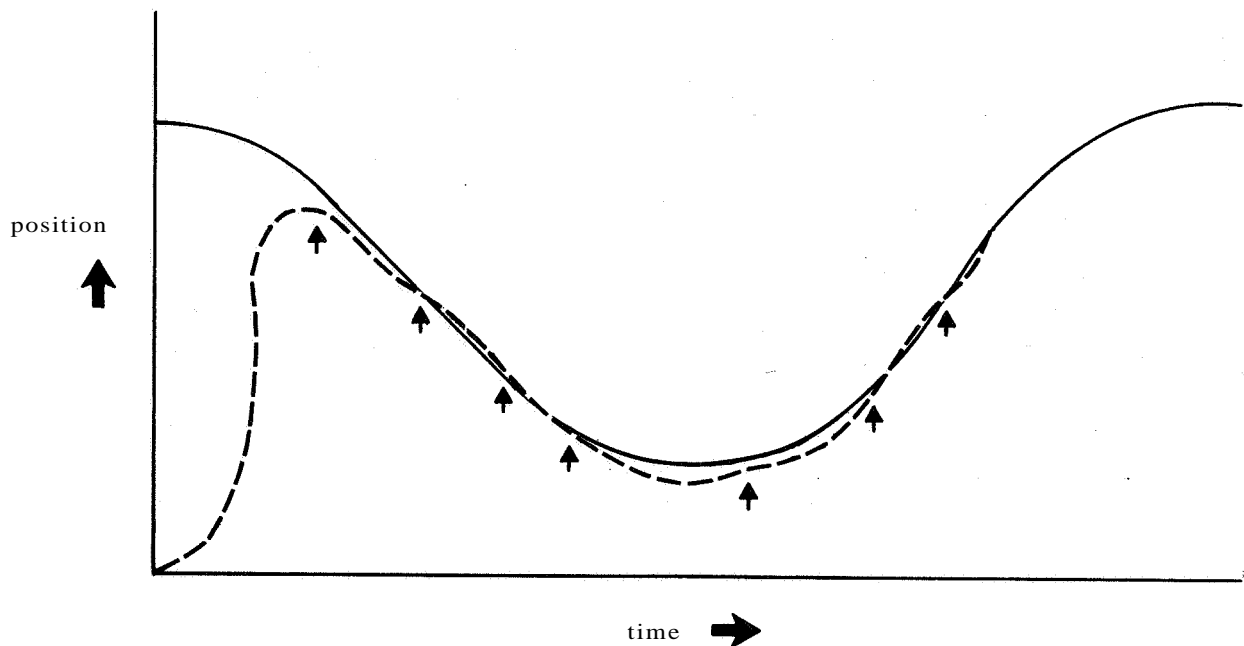


Figure 7.—Response to sine wave.

However, for the time-varying input, the situation is very much different. Here the particular table of combinations chosen is of prime importance. The problem is one of defining a suitable criterion function between error and error rate. For ideal tracking we would like to match velocities but since the corrections are of finite duration, for the time varying input, the time spent trying to match velocities might easily lead to large position errors.

An adjustment system which not only makes the thresholds more sensitive for improving results but also desensitizes the thresholds for poor results was found to be very unstable.

This, however, is probably due to the fact that a very small number of thresholds were used, two for error and one for error rate.

CONCLUSIONS

As pointed out previously, this project was only a feasibility study. No comparisons were made with actual human-operator data. The results are very encouraging. The sample model has many of the characteristics displayed by human operators.

The next step in the development of a more sophisticated model would be to include a small amount of memory. This would enable the model to respond much better to time-varying inputs.

Second, some work must be done on choosing the optimal set of initial parameters. Also, we would like to be able to choose the table of combinations by matching parameters with a real human operator.

Third, the duration of the corrections and possibly even the types of corrections can be made adaptive since the convergence rate is dependent on the duration of the correction. This might help explain some time-varying aspects of human operators.

REFERENCES

1. Elkind, J. I.; and Green, D. M.: Measurement of Time Varying and Nonlinear Dynamic Characteristics of Human Pilots. ASD tech. rep. 61-225, Dec. 1961.
2. Tomovic, R.; and McGhee, R. B.: A Finite-State Approach to the Synthesis of Bioengineering Systems. IEEE Trans. on Human Factors in Electronics, 1967.
3. Lemay, L. P.; and Westcott, J. H.: The Simulation of Human Operator Tracking Using an Intermittent Model. International Congress on Human Factors in Electronics (Long Beach, Calif.), May 1962.
4. Bekey, G. A.; and Angel, E. S.: Asynchronous Finite State Models of Manual Control Systems. Second Annual NASA-University Conference on Manual Control, NASA SP-128, 1966.

29. Obtaining Appropriate Human Pilot Describing Functions from Crossover Models and Optimal Control Theory

Lee Gregor Hojmann
Systems Technology, Inc.

N68-159301

Reference 1 presents the so-called crossover models for the human pilot as a function of the basic controlled element types, K_c , K_c/s , K_c/s^2 , and input bandwidth. My purpose here is to demonstrate one way in which optimal control theory may be employed to relieve what is often referred to as "artistry" in applying these crossover models to situations in which the controlled element is not clearly one of the basic types.

To accomplish this purpose I will perform the following:

- (1) Define a formulation of the problem to be solved
- (2) Indicate what is necessary from optimal control theory
- (3) Show a simple example: the pure gain controlled element
- (4) Show the results for a "complex" example: $Y_c = \frac{K_c}{s(s-1/T)}$

THE PROBLEM

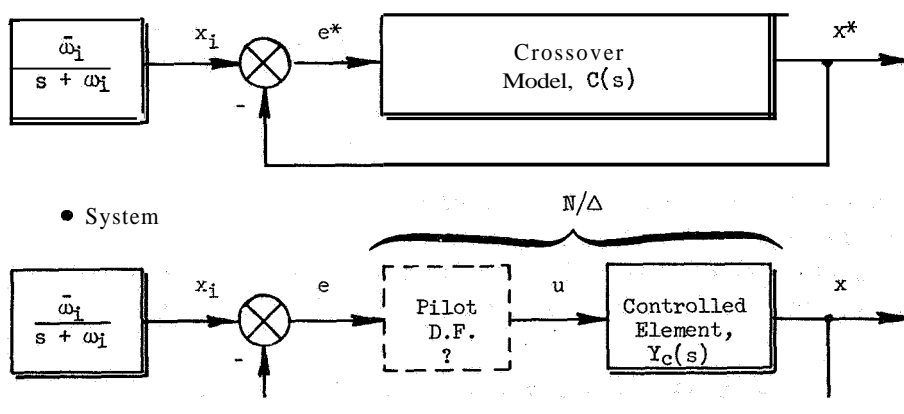


Figure 1.—The object of the problem is to minimize performance index WRT u where $PI = \int_0^\infty [q(x-x^*)^2 + r u^2] dt$. Note that transient analogs of stationary stochastic processes are used throughout.

*This research is part of that sponsored by the Air Force Flight Dynamics Laboratory, under Contract No. AF 33(615)-3652.

OPTIMAL CONTROL THEORY

For our problem let $(x - x^*) = \epsilon$, the response error. Then, in the frequency domain the performance index is

$$PI = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} (q \bar{\epsilon} \epsilon + r \bar{u} u) ds$$

The response error is

$$\epsilon(s) = Y_c(s) u(s) - C(s) e^*(s) = x(s) - x^*(s)$$

The first variation of **PI** on u must vanish at optimum, or

$$z(s) = (r + q \bar{Y}_c Y_c) u_0 - q \bar{Y}_c C e^*$$

$z(s)$ must be finite at each of its LHP poles.

The second variation of **PI** on u must be greater than zero, or the root square locus related quantity,

$$(r + q \bar{Y}_c Y_c) = 0$$

must not have closed-loop roots on Im axis.

We can find the optimal control using the so-called direct method presented in reference 2.

The form of optimal control is

$$u_0 = \frac{(a_n s^n + \dots + a_0) \bar{w}_i}{(\text{LHP zeros of } (r + q \bar{Y}_c Y_c) = 0) (\text{poles of } C e^*)}$$

The order of numerator is one less than denominator.

(1) Find a 's by substituting expression for u_0 into $z(s)$

(2) Require $z(s)$ to be finite at its LHP poles

(3) Gives algebraic equations for a 's

Having found the optimal control, it is an easy matter to develop Y_p using the fact that $Y_p Y_c$ is the open-loop transfer function of a unity-gain feedback loop.

SIMPLE EXAMPLE: $C(s) = \omega_c / s$, $Y_c = K_c$

Root square locus (trivial) is

$$r \left(1 + \frac{q K_c^2}{r} \right) > 0$$

$$z(s) = \frac{r \left(1 + \frac{qK_c^2}{r}\right) \bar{\omega}_1}{(s + \omega_1)} \cdot \frac{(a_1 s + a_0)}{(s + \omega_c)} \cdot \frac{-qK_c^2 \omega_c}{K_c}$$

$\frac{u_0}{x_1}(s)$

$a_1 = 0$

,

$a_0 = \frac{\omega_c}{K_c} \frac{\frac{qK_c^2}{r}}{1 + qK_c^2}$

But

$$\frac{x}{x_1} = \frac{u_0}{x_1} Y_c = \frac{N}{\Delta + N} = \frac{\frac{\omega_c}{K_c} \frac{\frac{qK_c^2}{r}}{1 + qK_c^2}}{s + \frac{\omega_c}{K_c} \frac{\frac{qK_c^2}{r}}{1 + qK_c^2}} = \frac{N}{s + \omega_c}$$

and

$$y_p = \frac{1}{Y_c} \frac{N}{A}$$

Given

$$Y_p = \frac{\frac{\omega_c}{K_c} \frac{\frac{qK_c^2}{r}}{1 + qK_c^2}}{s + \frac{\omega_c}{K_c} \frac{\frac{qK_c^2}{r}}{1 + qK_c^2}} = \frac{\omega_c / K_c}{s + \omega_c} \lim_{r \rightarrow 0}$$

COMPLEX EXAMPLE: $C(s) = \frac{-3.25 (s - 5.2)}{s (s + 5.2)} \frac{\omega_1 = 1.5}{r}$

$$Y_c = K_c / s (s - 1.0) \frac{qK_c^2}{r} = (10.)^4$$

Optimal controller :

$$y_p = \frac{-19.8}{K_c} \frac{s(s-1.0)(s-9.3)(s+11.3)}{(s-0.70)(s+5.0)[s^2+2(.54)10.7s+(1.07)^2]}$$

A comparison of open-loop functions is given in figure 2.

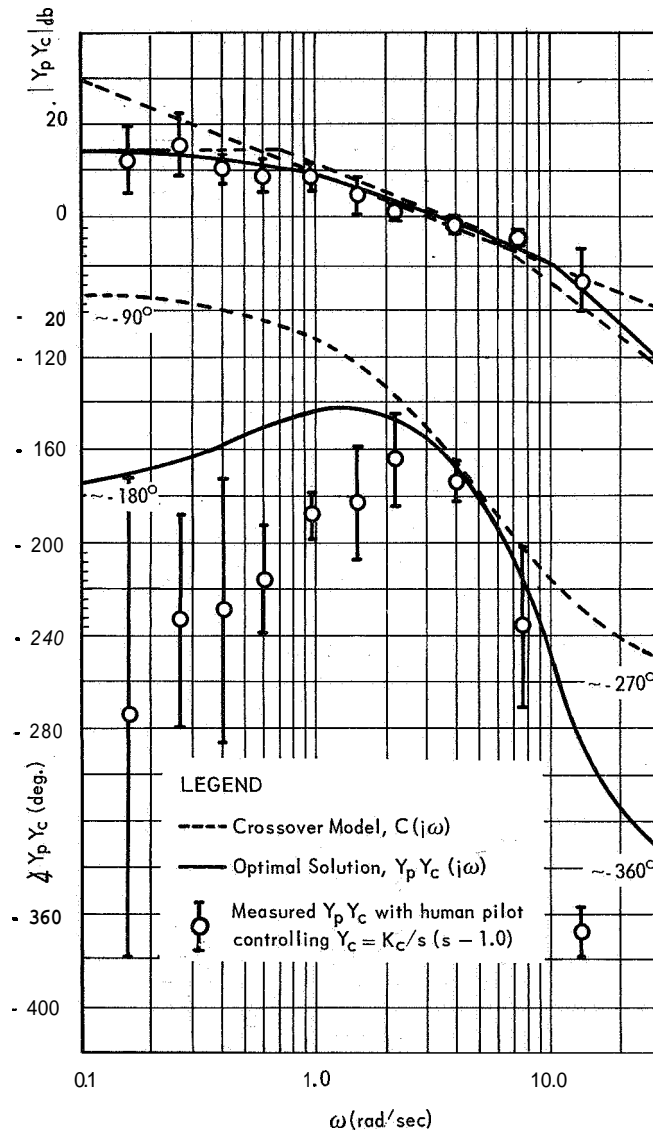


Figure 2.—Comparison of open-loop functions.

REFERENCES

1. McRuer, Duane; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems—Theory, Models, and Experiments With Controlled Elements and Forcing Function Variations. AFFDL-TR-65-15, Jan. 1965.
2. Rynaski, Edmund G.; and Whitbeck, Richard F.: The Theory and Application of Linear Optimal Control. AFFDL-TR-65-28, Jan. 1966.

X. APPLICATIONS

30. Application of Human Transfer Functions to a Design Problem

James J. Adams
Langley Research Center, NASA

N68-15931

An analytical design study was made of a proposed full-scale, manually controlled lunar-landing simulator using analytical transfer functions for the pilot-control response along with the analytical representation for the mechanisms. The simulator reproduced the lunar environment by supporting five-sixths of the weight of the test vehicle with an overhead cable. The cable was kept directly over the test vehicle by the automatic control of the longitudinal drive mechanism of the simulator. The results showed that the dynamic characteristics of the simulator that could be expected in the actual system were in a range that would influence the response of the manually controlled systems which were to be tested.

When the simulator was put in operation, the results of the analytical study were checked. The simulator was operated with the gain of the longitudinal drive set as high as was feasible with the actual mechanism and with a low gain to determine if this change would affect the pilot's response. The pilots reported that the degraded system was more difficult to control, and the records clearly showed a decrease in system damping with the degraded system.

One of the reasons for determining human transfer functions is to permit evaluation and prediction of the performance of manually controlled systems to be accomplished during design studies. This report explains how such transfer functions were used in the design analysis of the drive system of a lunar landing simulator.

The simulator was designed to provide a 400-foot by 50-foot by 180-foot-high volume in which lunar-landing maneuvers could be studied. The lunar gravity was simulated by supporting five-sixths of the weight of the test vehicle by a cable. The load in the cable was regulated by measuring the load with a strain gage and operating the overhead winch in response to the error in this measured load. The cable was kept directly over the vehicle by measuring the cable angle at the overhead, traveling bridge and by moving the bridge in response to this measured angle. It is this longitudinal drive system of the bridge that is the subject of the design study reported in this paper.

The transfer functions used to describe the pilot's control action were derived in reference 1. These transfer functions describe the pilot's control used when controlling the multi-loop system representative of the lunar-landing horizontal translation maneuver. The use of these pilot transfer functions in determining the most suitable simulator drive characteristics will be presented.

SYMBOLS

| | |
|---|---------------------|
| m | mass, slugs |
| x | translation, ft (m) |

| | |
|----------------------------------|---|
| l | pendulum length, ft (m) |
| φ | pendulum angle, deg |
| g | gravity, 32.2 ft/sec ² (9.81 m/sec ²) |
| $K_{\varphi}, K_{\dot{\varphi}}$ | simulator drive system control gains |
| s | Laplace operator, per sec |
| θ | pitch attitude angle, deg |
| δ | control moment, rad/sec ² |
| ω | undamped natural frequency, rad/sec |
| h | maximum deflection of first cable vibration mode, ft (m) |
| K_1, K_2, τ | gains in analytical transfer function of pilot |
| Q_L | leakage flow, in. ³ /sec |
| Q_{RL} | relief valve flow, in. ³ /sec |
| V | oil volume under compression, in. ³ |
| B | bulk modulus of oil, psi |
| P_m | motor pressure, psi |
| D_m | motor volume, in. ³ |
| J_m | motor inertia, in.-lb/sec ² |
| w_m | motor rotation, rad/sec |
| n | gear ratio |
| B_m | motor damping, in.-lb/rad/sec |
| K_L | tire torsional spring constant, in.-lb/rad |
| T_L | tire torque, lb |
| J_L | load inertia, in.-lb-sec ² |
| B_L | load viscous friction, in.-lb/rad/sec |

Subscripts:

| | |
|---|---------|
| B | bridge |
| V | vehicle |
| C | command |
| e | error |

DESCRIPTION OF PROBLEM

SIMULATOR.—Simplified equations of motion for the longitudinal motion of the overhead bridge and pendulum consisting of the cable-supported vehicle are

$$\begin{aligned} (m_B + m_V) \ddot{x}_B + m_V l \ddot{\phi} &= 0 \\ m_V l \ddot{x}_B + m_V l^2 \ddot{\phi} + m_V g l \phi &= 0 \end{aligned}$$

The characteristic equation for the system in Laplace rotation is

$$(m_B + m_V) (m_V l^2) s^2 + (m_B + m_V) (m_V g l) - m_V^2 l^2 s^2 = 0$$

which is the familiar equation for a pendulum in which the frequency is primarily determined by the length l but with an additional term representing the influence of the bridge being free to move, which reduces the frequency somewhat.

If the control commands that the bridge accelerates as a function of ϕ and $\dot{\phi}$, the equations of motion become

$$\begin{aligned} (m_B + m_V) \ddot{x}_B + m_V l \ddot{\phi} &= -K_\phi \phi - K_{\dot{\phi}} \dot{\phi} \\ m_V l \ddot{x}_B + m_V l^2 \ddot{\phi} + m_V g l \phi &= 0 \end{aligned}$$

The characteristic equation, in Laplace rotation, is

$$(m_B + m_V) (m_V l^2) s^2 + (m_V l K_\phi) s + (m_B + m_V) (m_V g l) - m_V^2 l^2 s^2 + m_V l K_{\dot{\phi}} s = 0$$

It can be seen from this equation that the K_ϕ gain, in the coefficient of s , of the drive-system control will supply damping to the system, and the $K_{\dot{\phi}}$ gain will increase the pendulum frequency to keep the bridge above the suspended vehicle.

The bridge drive unit was an electrohydraulic unit consisting of a synchronous electric motor which drove a variable-displacement hydraulic pump. Control of the bridge was exercised by the operation of a pump stoker, which controlled the displacement of this pump. The pump drove the fixed-displacement hydraulic motors attached to the wheels of the bridge. A fixed displacement of the stoker produced a steady-state constant velocity of the bridge.

Since the stoker controlled bridge velocity instead of bridge acceleration, the control function of accelerating the bridge as a function of pendulum angle was achieved by displacing

the stoker as a function of the integral of the pendulum angle; the function of accelerating the bridge as a function of rate of change of pendulum angle was achieved by displacing the stoker as a function of pendulum angle.

The predominant dynamic characteristic of the drive unit was the oscillatory response of bridge velocity to stoker displacement that resulted from the compressibility of the hydraulic fluid as it reacted against the mass of the bridge. This dynamic characteristic is expressed by the equation.

$$\frac{\dot{x}_B}{\text{Stoker displacement}} = \frac{\frac{c_1}{d_m}}{\frac{B}{V} \frac{J_L}{d_m} s^2 + \frac{J_L K_L}{d_m} s + 1}$$

where

c_1 oil volume delivered by pump per unit displacement of stoker, in.³/unit

d_m motor displacement, 28 in.³/rad

B bulk modulus of oil, 1×10^5 psi

V oil volume under compression, 600 in.³

J_L total inertia of load reflected at motor output shaft, 3050 in.-lb/sec²

K_L leakage coefficient, in.³/sec/psi

For the system under study, the natural frequency of the drive unit as determined by this equation is 4.6 rad/sec. The detailed analysis of the system naturally included this characteristic.

In addition to the drive unit dynamics, the following dynamic factors were also included in the detailed analysis:

- (1) The compliance of the primary electric motor
- (2) The time constant and the limit displacement of the stoker
- (3) The motor leakage as a function of $\sqrt{\text{Pressure}}$
- (4) Pressure relief valve flow
- (5) Nonlinear friction of the gearing
- (6) Tire compliance

Another important dynamic characteristic of the system was the oscillatory characteristic of the cable. Cable vibrations added to the measured cable angle and therefore added a spurious signal to the control signal. Equations for the first two modes of vibration for different, fixed cable length were determined and used in the analysis. These equations included the effect of a lumped mass located near the vehicle that represented the whiffletree which was a part of the support and gimbal arrangement of the vehicle. The frequency of the first mode for a 200-foot cable length was 8.72 rad/sec, which is very close to the natural frequency of the drive unit, and which therefore put a limit on the precision with which the bridge could be maintained above the vehicle. This vibration frequency would increase at shorter cable

lengths and also change with vehicle weight. The analysis was made for fixed cable lengths and vehicle weights.

PILOT TRANSFER FUNCTION.—The pilot transfer functions, which were used in conjunction with the simulator equations, are derived in reference 1 and are repeated here. The pilot-vehicle system involved in the landing maneuver is a multiloop system described in the block diagram presented in figure 1. The inner loop deals with the attitude control of the vehicle. The vehicle response to attitude control was assumed to contain a proportional rate feedback and is given by the equation

$$\frac{\theta}{\delta} = \frac{0.5}{s(s+0.5)}$$

which defines a rate system with a rate response having a time constant of 2 seconds. Reference 1 demonstrates that a pilot's response in such an inner loop is given by

$$\frac{\delta}{\theta_{c,e}} = \frac{K_1 \tau \left(1 + \frac{K_2}{\tau} s \right)}{(s + \tau)^2} = \frac{96(1 + 0.4s)}{(s + 6)^2}$$

The combination of the pilot and vehicle gives this inner loop a closed-loop characteristic frequency of 1.2 rad/sec and a damping ratio of 0.26.

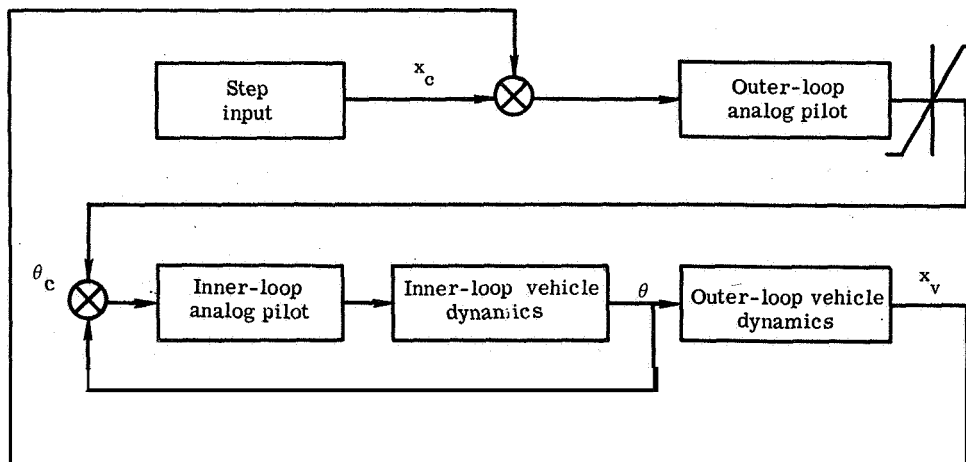


Figure 1.—Block diagram of the double control loop.

The outer loop of the system deals with the longitudinal translation, and the vehicle response to attitude angle is given by a pure inertial response

$$\frac{x}{\theta} = \frac{5.36}{s^2}$$

This relation is derived from the linearized equation of motion for the horizontal component of acceleration due to one-sixth of thrust that would be in effect in the lunar environment

$$\ddot{x} = -\frac{1}{6} g \sin \theta$$

using small-angle linearization

$$\ddot{x} = \frac{1}{6}g\theta = 5.36\theta \text{ ft/sec}^2 = 1.636 \text{ m/sec}^2$$

$$\frac{\ddot{x}}{\theta} = 5.36 \frac{\text{ft/sec}^2}{\text{rad}} = 1.63 \frac{\text{m/sec}^2}{\text{rad}}$$

Reference 1 demonstrates that the pilot response in such an outer loop is

$$\frac{\theta}{x_e} = \frac{0.9(1+9.2s)}{(s+10)^2}$$

This pilot response defines a characteristic response of the complete system which has two small real roots, $s = -0.167$ and $s = 0.336$. In terms of an oscillatory response, these two roots define an overdamped response with a natural frequency given by

$$\omega = \sqrt{(0.167)(0.336)} = 0.236 \text{ rad/sec}$$

This system frequency characterizes the translation response of the system. Since the response characteristic of the longitudinal drive system of the simulator must have a response frequency higher than that of the system which is to be tested, this calculated response characteristic of the pilot-controlled translation response provides a first, rough criterion for the required characteristics of the longitudinal drive system.

ANALYSIS AND TEST RESULTS

A detailed analytical study was made to determine precisely the drive system characteristics and the suitability of the characteristics. The computer diagram used in this study is presented in figure 2. It includes the complete representation of the drive system, the cable dynamics, and the pilot controlled vehicle representation.

The first phase of the study was conducted to determine just how fast the bridge could be made to respond. Open-loop step thrust inputs to the vehicle were used as the forcing function in these studies. The results showed that the presence of the cable vibration modes of motion in the system placed an upper limit on the pendulum damping gain K_{ϕ} . If this gain was adjusted too high, the first vibration mode would become unstable, as is illustrated in figure 3.

The limit on the pendulum damping gain placed further restriction on the pendulum frequency gain K_{ϕ} . The range of possible system characteristics that could be achieved is shown in figure 4, which shows the vehicle velocity response to a 2-second thrust impulse. The oscillatory nature of these responses is the result of the bridge drive system characteristics. It can be seen that a well-damped response with a frequency of 1.57 radians/sec (a period of 4 sec), or a poorly damped response with a frequency of 2.5 rad/sec (a period of 2.5 sec) could be achieved.

Both of these frequencies are above the 0.236-rad/sec frequency for the pilot controlled translation response of the lunar landing system. However, it cannot be confidently concluded that they are sufficiently high to have no effect on the simulation. To determine what effect the bridge response might have on the pilot-controlled maneuver, the analytical representation of the pilot and vehicle were included in a closed-loop representation of the complete system, and the response to a commanded 200-foot displacement was determined. The

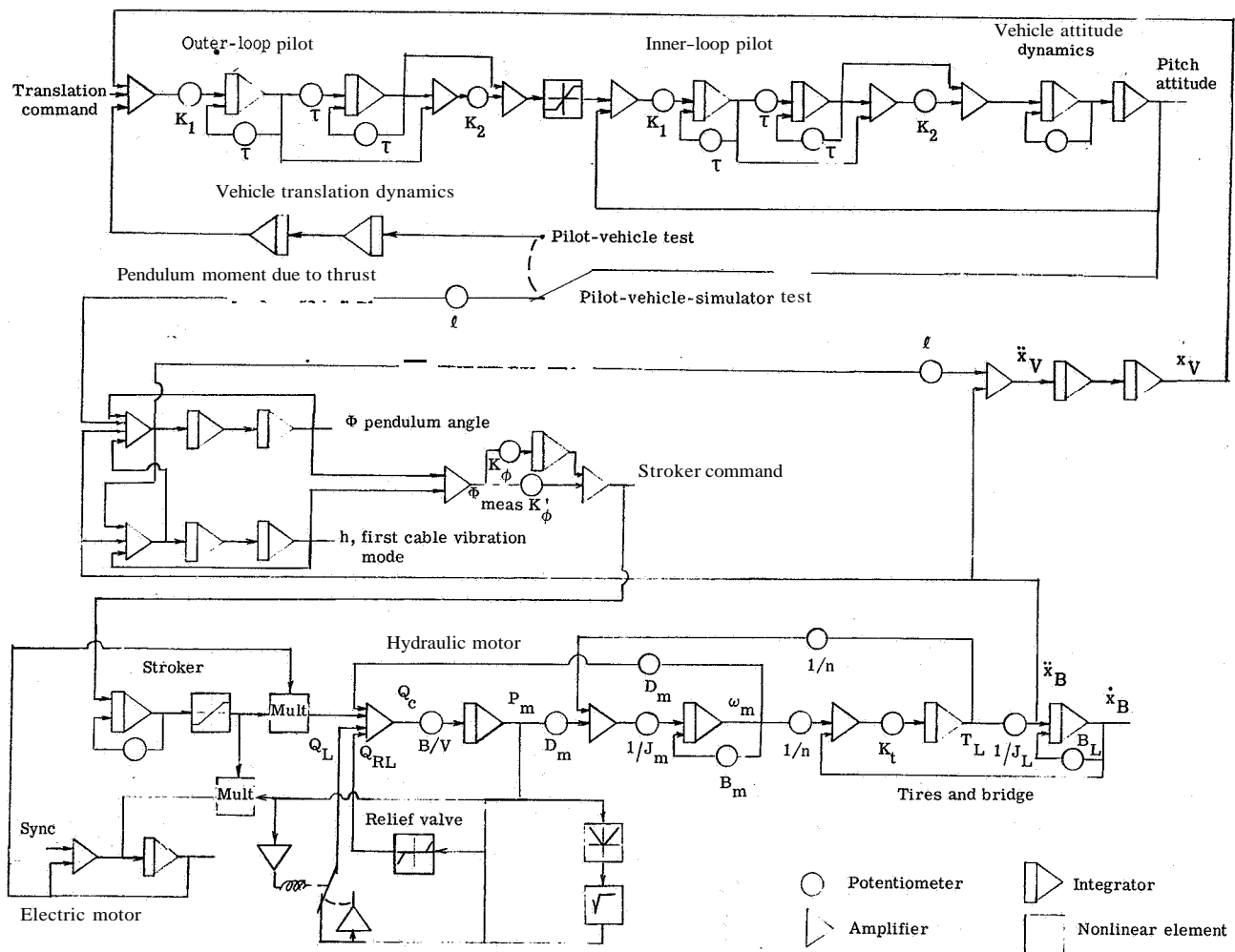


Figure 2.—Diagram of pilot-vehicle-simulator system.

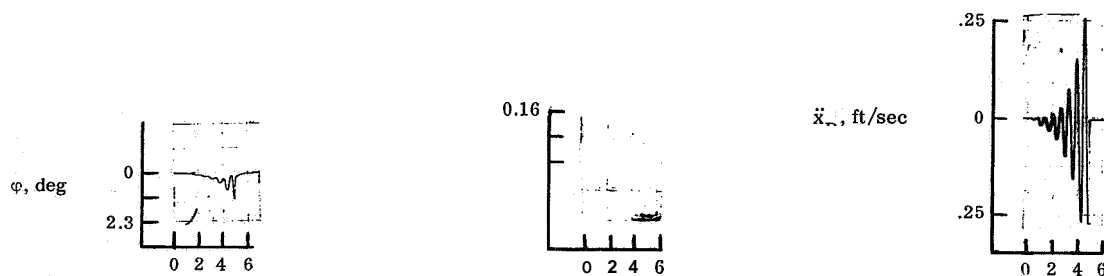


Figure 3.—Calculated cable divergence.

results are presented in figure 5, which show, first, the response of the pilot-vehicle combination alone, and then the response of the pilot-vehicle-simulator combination with the two different simulator characteristics presented before. Using the pilot-vehicle combination response as the standard for comparison, it can be seen that including the simulator bridge dynamics in the loop does indeed influence the response. With the lower gain bridge control the system is degraded to the point of instability. It was therefore concluded that the simulator should be adjusted to have as high a frequency characteristic as possible. Also, it was

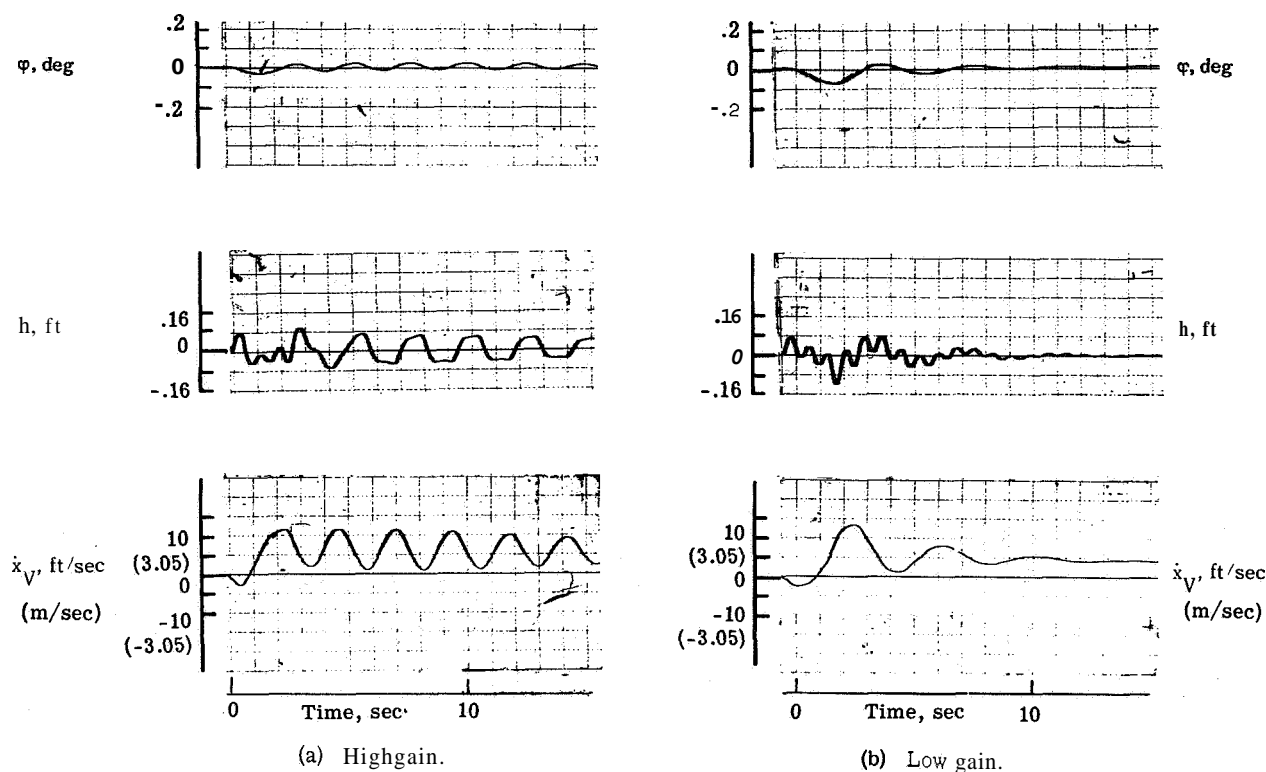
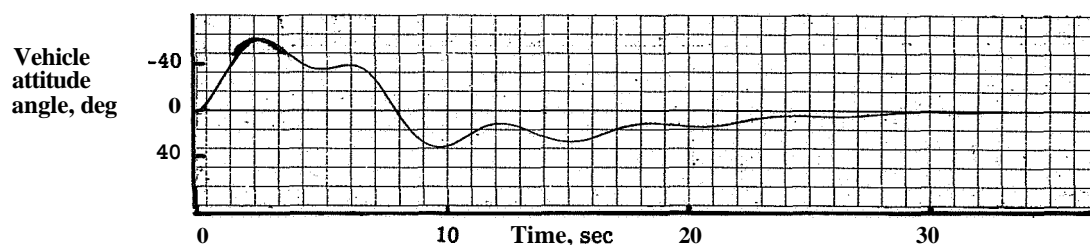


Figure 4.—Calculated response of simulator to 2-second thrust impulse.

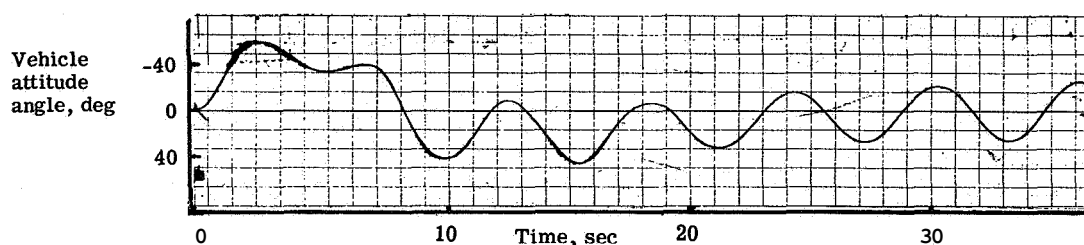
indicated by the analysis that the pilots might find the simulator slightly more difficult to control than the real lunar-landing system.

When the simulator was put in operation, the system characteristics which could be achieved with the actual mechanism were determined. It was found that the highest stable pendulum frequency that could be obtained was 1.4 rad/sec (a period of 4.5 sec). Piloted runs were then made with this highest frequency response of the bridge, and with the K_ϕ gain placed at a lower setting to check the analytical results to determine how such a change would affect the piloted system response. The lower frequency used was approximately 0.8 rad/sec (a period of 8 sec). Figure 6 shows cable angle responses of the simulator in these two conditions to open-loop step impulses. The piloted tests were started with the vehicle hovering at an altitude of approximately 30 feet. The pilot then translated the vehicle 200 feet and attempted to stop and hover over a mark located on the ground. The run with the higher response was made first, and then immediately a second run was made with the lower setting. Two different pilots were used.

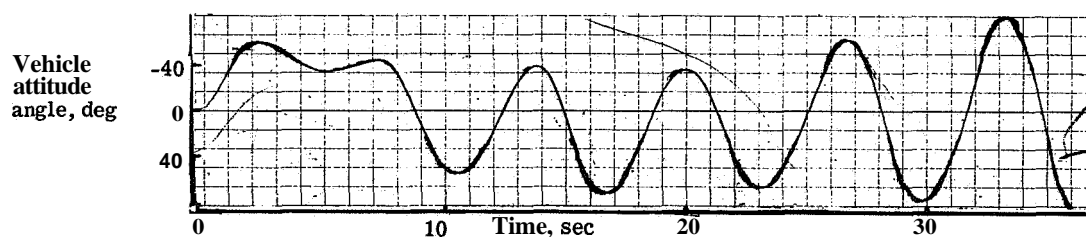
Sample time histories of these tests are shown in figure 7. In this test the first maneuver, from the 50-foot point to the 275-foot point, was done with the high-gain drive system. At the 50-second mark, while the pilot was turning 180°, the drive system gain was readjusted to the low-gain setting, and the pilot started to go back to the 50-foot mark. It can be seen that the general nature of these maneuvers is very similar to that computed in the analytical study. With the lower response characteristics for the simulator a very noticeable decrease in damping of the attitude angle can be seen. The pilot did not continue the maneuver in this case, but rather dropped the intention to control translation precisely. He stopped the attitude oscillation and landed at the point that was below him at that time. The pilots commented that in the lower response runs they were having more difficulty in controlling the vehicle, and as one pilot said, he felt he was in a "pilot-induced oscillation" condition.



(a) Pilot vehicle combination.



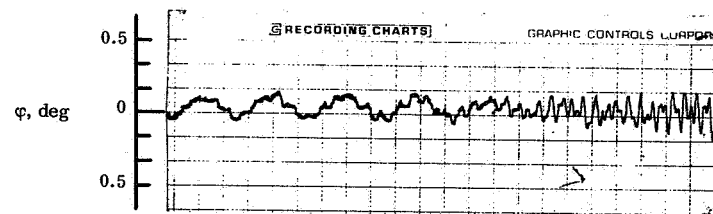
(b) Pilot-vehicle-simulator combination with high gain simulator.



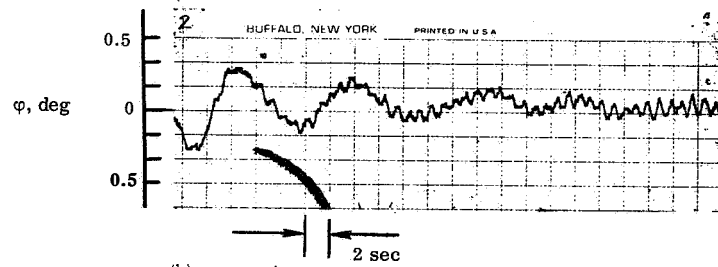
(c) Pilot-vehicle-simulator combination with low gain simulator.

Figure 5.—Calculated vehicle attitude response to a 200-foot translation command.

Since the simulator characteristics that were achieved with the actual mechanism when it was put in operation were not the same as those determined in the analytical study, and since the pilot's response showed a lower attitude angle limit in the flight tests than was assumed in the analytical study, the analytical study was repeated in an attempt to reproduce more closely the flight time histories. In these repeated calculations the drive system gains were adjusted to give a poorly damped 4.5-second pendulum period in one case and a well damped 8-second pendulum period in the second case to correspond to the two conditions that were tested in the flight tests. The same linear transfer functions were used for the representation of the pilot. The attitude limit was placed at 10° , which corresponds more closely with the limit used by the pilot in the flight tests than did the 40° limit used in the initial analytical study. With this 10° limit on attitude angle, a well-controlled attitude time history was calculated, shown in figure 8(a), which is a better reproduction of the flight time history than was the initially calculated response. When the simulator characteristics were changed so as to have an 8-second period, a deterioration in the stability of the calculated attitude-angle



(a) High gain.



(b) Low gain.

Figure 6.—Simulator cable angle response to small inputs.

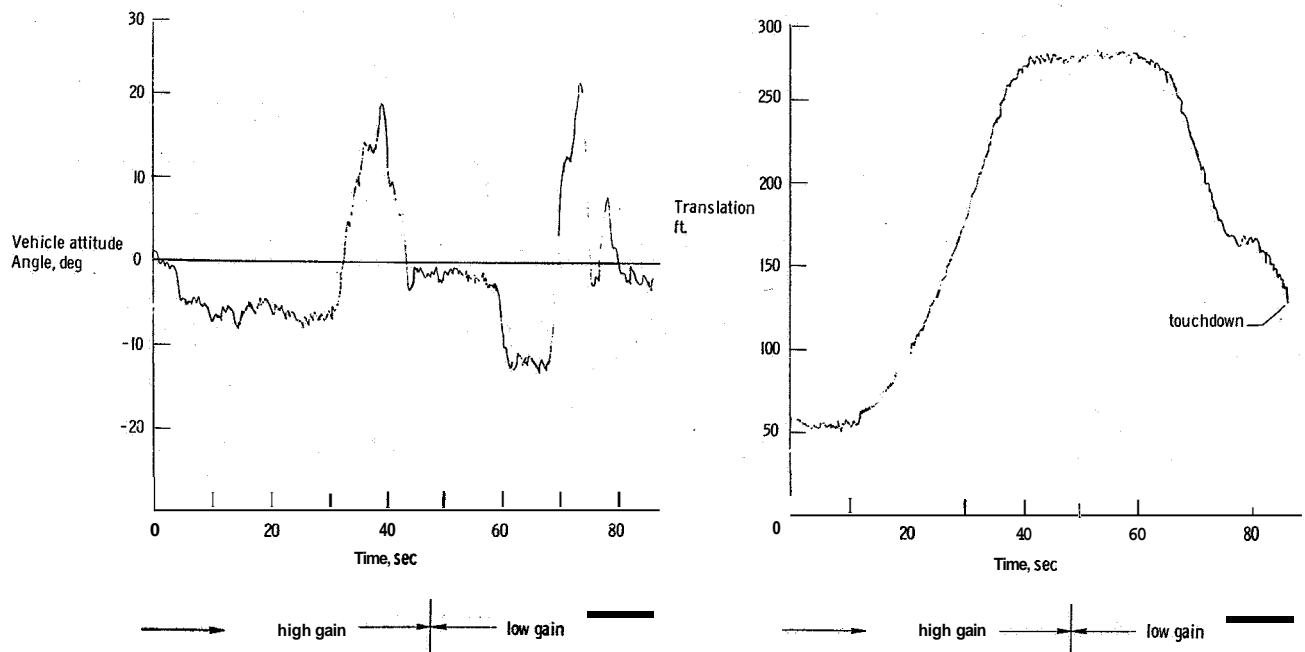


Figure 7.—Simulator response to a 200-foot translation command with both high-gain and low-gain simulator control.

time history resulted (fig. 8(b)). These repeated calculations further confirm the conclusion drawn from the initial analytical study.

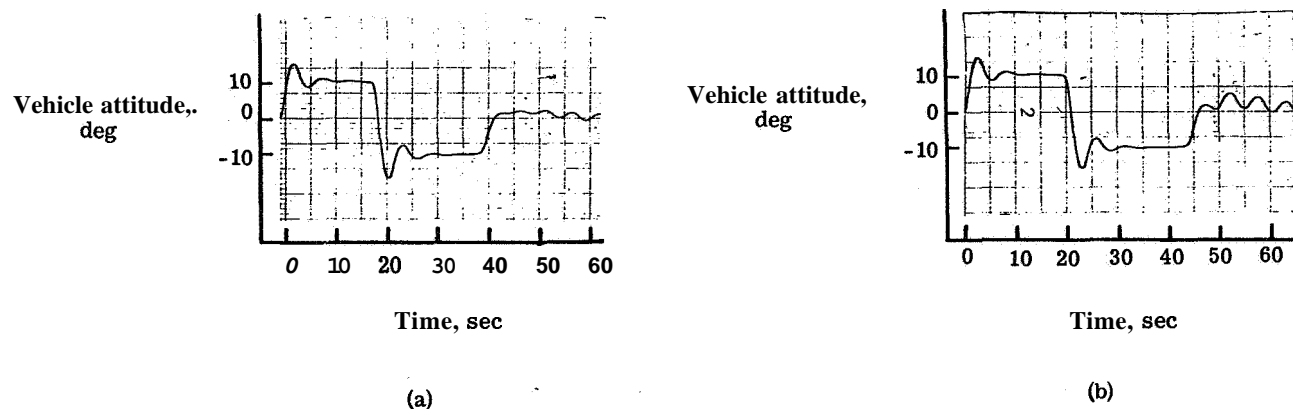


Figure 8.—Repeated calculation of vehicle attitude response to a 200-foot step translation command. (a) Pilot-vehicle-simulator combination with low gain (4sec period) simulator; (b) pilot-vehicle-simulator combination with very low gain (8sec period) simulator.

CONCLUSIONS

The experience gained in the exercise of predicting the characteristics of a manually controlled simulator and verifying these characteristics when the simulator was put in operation has demonstrated the validity and usefulness of analytical expression of human response. The analysis showed that the range of simulator dynamics which was likely to occur in the simulator would influence the response of the pilot-controlled maneuver, and tests with the completed hardware confirmed this conclusion.

It was concluded that the gain of the simulator longitudinal drive system should be kept as high as possible to minimize the effect on the piloted maneuvers, and it was indicated by the analysis that the tasks performed with the simulator might be slightly more difficult than the same tasks performed in the lunar environment.

REFERENCE

1. Adams, James J.; Bergeron, Hugh P.; and Hurt, George J.: Human Transfer Functions in Multi-Axis and Multi-Loop Control Systems. NASA TN D-3305, 1966.

31. Asymptotic Stability Studies in Simulated Car Following'

Robert E. Fenton
Ohio State University

N68-15932

A control stick with a built-in tactile aiding device was tested in a simulated car-following situation. The tactile device, or "finger," gave the driver of a following car information—headway and relative velocity—about the state of a lead car. The finger-position sensitivity and the range of tactile aiding ratios under which a driver may minimize headway and relative velocity variations are obtained, together with a mathematical model of the display-driver-control stick combination. The corresponding driver-vehicle system is shown to be locally, but not asymptotically, stable. Hence, optimum tracking and asymptotic stability do not occur under identical conditions.

Asymptotic stability may probably be achieved by compensating the driver-vehicle system. To this end, various compensation methods are examined, and models are obtained for the modified system. Insight into the form of the required compensation was gained; however, all of the systems examined were asymptotically unstable.

Traffic flow on crowded highways may be increased by reducing the spacing or headway between adjacent vehicles. This may be safely done either by using an automatic longitudinal control system in individual vehicles, or by improving the driver-vehicle system. The second approach is discussed in this paper; namely, the use of a modified manual control unit for improved car-following performance.

A number of investigations have shown that the headway in steady-state car following is not invariant. For example, Bierley (ref. 1) found that under nearly ideal test track conditions, the driver of a following car was unable to maintain a constant headway. Additionally, in a recent study of car following under normal roadway conditions, Montano (ref. 2) found that sizable headway variances existed. Since a time-varying headway decreases the efficiency of traffic flow, it is desirable to overcome or reduce such variation.

There is considerable evidence which indicates that this variation is due to an inherent limitation of the driver—namely, his inability to detect slow relative motion between two vehicles. This limitation may be overcome by presenting additional headway information to the driver. Bierley (ref. 1), in a study of intervehicle spacing displays, found that considerable reduction in headway variation could be achieved by using a galvanometer which indicated the instantaneous spacing error. In a similar study, Rockwell and Gantzer (ref. 3) found that the use of a visual display greatly reduced headway and velocity variance at large headways; however, at small headways the drivers would not use the display. Therefore, it appears that the use of a visual display is not promising.

*The research reported in this paper was sponsored by the Ohio Department of Highways in cooperation with the U. S. Department of Public Roads.

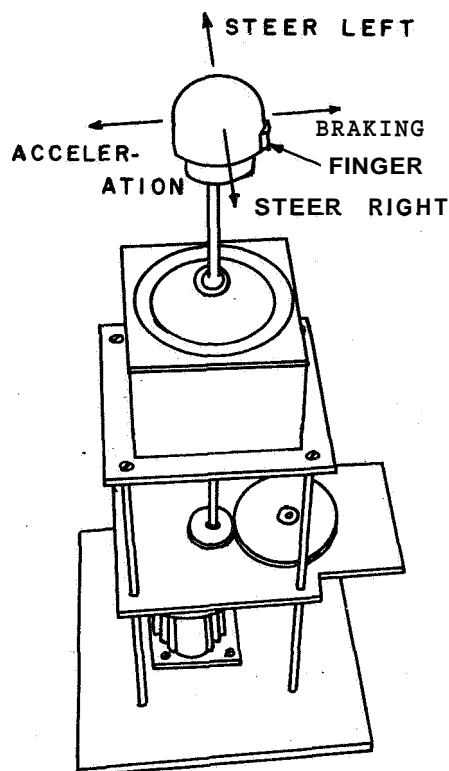


Figure 1.—Control stick with finger.

Fenton (ref. 4) has studied a kinesthetic-tactile display, in which headway information was imparted to the driver via a finger which was mounted in the head of a control stick (see fig. 1). This stick was used instead of conventional automobile controls. The position of the finger was servo-controlled and gave the driver an indication of his instantaneous deviation from some desired headway. An automobile simulator was used for this study.

Sizable reductions in both headway and velocity variance were obtained when the finger position was partially quickened; that is, excited by a single proportional to a linear combination of headway deviation and relative velocity. These reductions in variance were obtained for both small and large headways. Montana (ref. 2) obtained similar results from a real-world study of this tactile display. Hence, it appears that a tactile display does not have the same limitation as a visual display.

However, improved tracking performance is not the only criterion by which the use of the display should be judged. It is also necessary to insure that the corresponding driver-vehicle system is both locally and asymptotically (long-line) stable.

A number of investigators have obtained models for the conventional driver-vehicle system. The most widely publicized such model—the non-linear Herman equation (ref. 5)—has been used by Nemeth and Reebel (ref. 6) to show that the conventional system is asymptotically unstable. The same conclusion may be drawn from a study of the linear model suggested by Hanken (ref. 7).

Barbarosa (ref. 8) has shown that a disturbance will be attenuated as it propagates down a line of traffic if

$$\left| \frac{V_n(j\omega)}{V_{n-1}(j\omega)} \right| \leq 1 \quad \text{for all } \omega \quad (1)$$

where:

$V_n(j\omega)$ Fourier transform of the velocity of the n th car in the queue

$V_{n-1}(j\omega)$ Fourier transform of the velocity of the $(n-1)$ car

This equation, which is valid for the small-signal case, may be used to determine if a system is asymptotically stable, provided a linear representation may be obtained for that system.

A block diagram of the driver-vehicle system is shown in figure 2. It is assumed that the driver is primarily concerned with longitudinal control and that only small lateral corrections need be made. (This corresponds to steady-state car following on a long, straight superhighway.) Note the presence of two inputs to the driver—his normal visual input and one from a display. If the driver receives sufficient information from the display to control

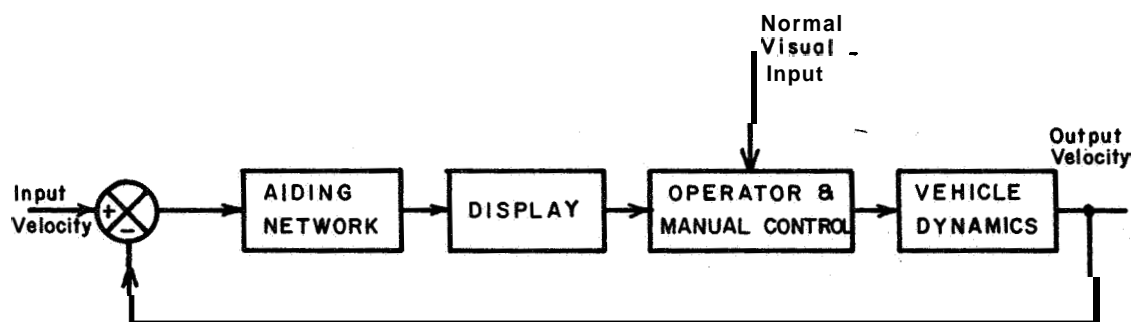


Figure 2.—Block diagram of man-vehicle system.

the vehicle, the display-driver-control combination may be treated as a single input, single output system. A linear model $Y_p(j\omega)$ may then be obtained by using well-known techniques of time-series analysis.

The closed-loop transfer function of the driver-vehicle system is

$$\frac{V_2(j\omega)}{V_1(j\omega)} = \frac{Y_1 Y_2 Y_p}{1 + Y_1 Y_2 Y_p} \quad (2)$$

where

- $V_2(j\omega)$ velocity of the driven car
- $V_1(j\omega)$ velocity of a lead car
- $Y_1(j\omega)$ transfer function of aiding network
- $Y_2(j\omega)$ transfer function of vehicle longitudinal dynamics

Here, the effects of operator-generated noncoherent noise have been neglected. In order to obtain asymptotic stability, it is necessary that

$$\left| \frac{Y_1 Y_2 Y_p}{1 + Y_1 Y_2 Y_p} \right| \leq 1 \quad \text{for all } \omega \quad (3)$$

Letting

$$Y_1(j\omega) Y_2(j\omega) Y_p(j\omega) = Y \cos \alpha + Yj \sin \alpha$$

and squaring, one has

$$Y^2 \leq 1 + Y^2 + 2Y \cos \alpha$$

or

$$0 \leq 1 + 2Y \cos \alpha \quad (4)$$

Two necessary conditions may be derived from equation (4); first, the phase margin

$$\psi \geq 60^\circ$$

and secondly, the gain margin

$$\text{GM} \geq 6 \text{ db}$$

The condition for local stability is that these quantities are greater than zero.

An extension of previous work on the tactile controller is presented in this paper. The conditions under which optimum car-following tracking performance is obtained are specified, together with an asymptotic stability study of the driver-tactile controller system. All results reported here were obtained using an automobile simulator.

OPTIMUM TRACKING CONDITIONS

The greatest improvement in car-following performance was previously obtained when the finger position x was proportional to a linear combination of the headway deviation h and the relative velocity v .

$$x = K_1 h + K_2 v$$

The problem of determining the conditions under which optimum car-following tracking performance is obtained may be reduced to one of determining the best combination of finger-position sensitivity and aiding ratio K_2/K_1 . (Finger position sensitivity refers to the finger movement per unit headway deviation.) To this end, a set of car-following experiments were conducted using an automobile simulator.

THE AUTOMOBILE SIMULATOR. All testing was performed with the automobile simulator shown in figure 3 (ref. 9). This simulator was designed to simulate a two-car situation in which a lead vehicle is closely followed by a second vehicle. The subject, or

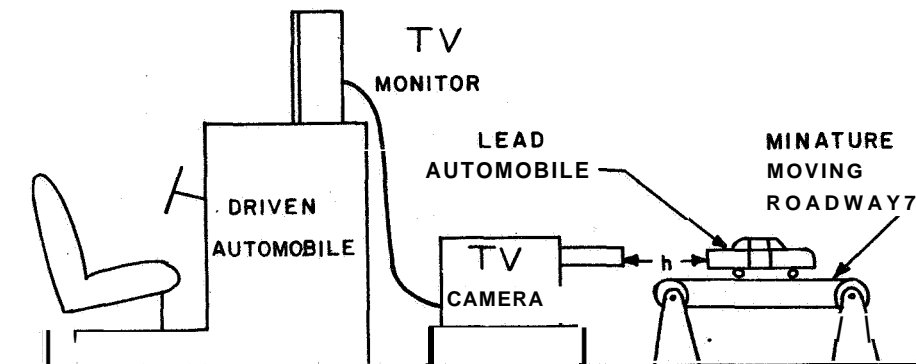


Figure 3.—The automobile simulator.

"driver" of the rear vehicle, is able to exercise control over the relative velocity and headway between the two cars. He is also able to steer; however, the following vehicle is confined to one lane; that is, no passing is allowed.

A picture of the lead car and roadway is shown on the television screen. Such a picture is intended to approximate the complex spatial-time display which a driver sees under daylight driving conditions. Additional input data were obtainable from a speedometer mounted on the "dash" in front of the subject. An analog computer was included to simulate the characteristics of the following car and automatically perform the necessary calculations.

EXPERIMENTAL DESCRIPTION. The lead-car velocity consisted of a constant value plus low-frequency random noise, the noise being such that the lead-car accelerations were in the range

$$-0.03 \leq a \leq 0.03g$$

The corresponding velocity variance of the lead car was $3 \pm 0.5 \text{ (ft/sec)}^2$. The subject was instructed to maintain a constant headway in the presence of the noise. The headway in feet was to correspond to the average speed on a 1 to 1 basis—1 foot for each mile per hour of speed. This spacing is smaller than that recommended by various state highway departments; however, since one goal of this research is improved methods for reducing headway, it seems desirable to test all proposed systems at such decreased values of headway.

The subject was also given a second task: a steering task in which he had to overcome the effects of random noise which caused apparent lateral motion of the driven car. This noise was essentially uncorrelated with the velocity noise as may be seen from the spectra shown in figure 4.

Three male subjects from 20 to 23 years old were used in this experiment. These subjects all had driving experience and had either 20/20 vision or vision corrected to 20/20.

Three values of sensitivity, 0.00185 in./ft, 0.053 in./ft, and 0.107 in./ft, hereafter referred to as low, medium, and high, were examined, together with four values of aiding

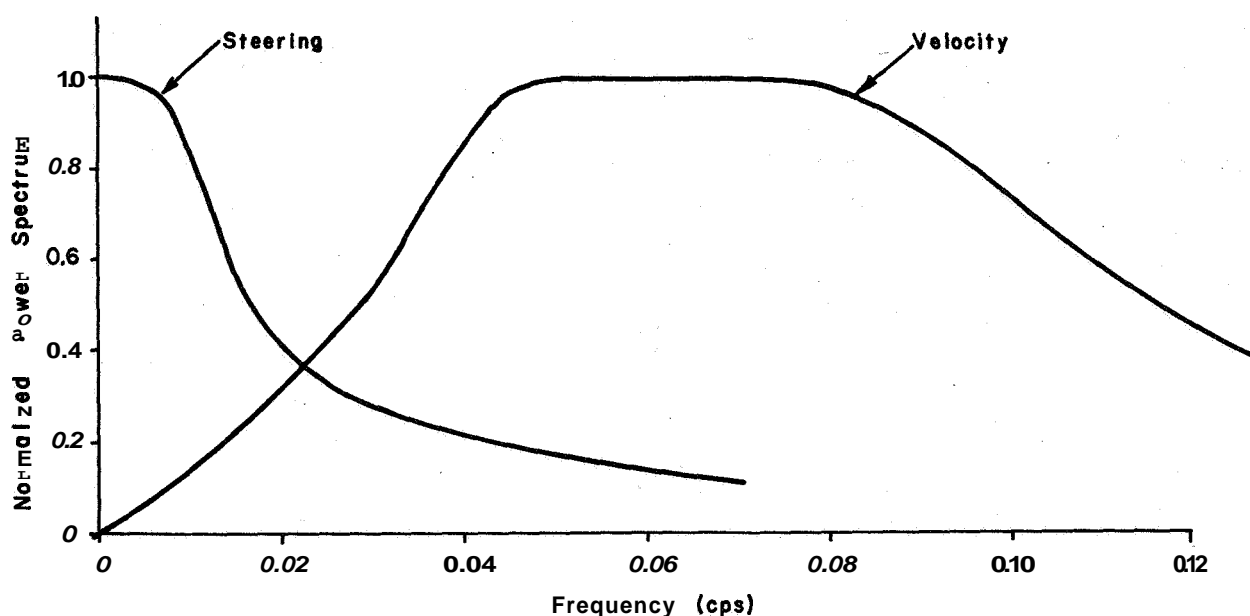


Figure 4.—Power spectrum of velocity and steering noise.

ratio—1.6, 3.2, 6.4, and 12.7. The experimental design was not orthogonal, for it quickly became evident that superior results were obtained with medium sensitivity. Hence, the low and high sensitivities were only tested with aiding ratios of 1.6 and 3.2, while the medium sensitivity case was tested with all ratios.

Each of the above mentioned combinations was tested at three values of lead-car average speed, 40, 50, and 70 miles per hour. Each subject made three 12-minute runs at each of these speeds with each combination of sensitivity and aiding ratio.

EXPERIMENTAL PROCEDURE. Before any data were collected, the subjects were given time to "drive" the simulator and thus become familiar with the combination being tested. Two half-hour familiarization periods for each type sufficed, it being assumed that this pretesting removed the learning effects of subjects from the final experimental results.

For a given type of control and a given average speed of the lead car, say 40 mph, the subject was first shown the lead car at a simulated headway of 40 feet and told to maintain that headway. The subject was then given control of the relative state of the two cars. Some two minutes after he assumed control, the variance measuring circuits were connected, and headway and velocity variance determined Over the succeeding 12-minute period. After such a run, the subject was given a 5-minute rest and then tested for 12 more minutes. This process continued until the subject completed three 12-minute runs, a different lead-car average velocity being used for each run.

EXPERIMENTAL RESULTS. The headway variances obtained from the three subjects for a given sensitivity and aiding ratio—some nine values—were averaged to give a single value (the individual data from this experiment are contained in ref. 10). The average values for various sensitivities and an aiding ratio of 1.6 are plotted versus lead-car average speed in figure 5. The corresponding average velocity variances are similarly plotted in figure 6.

An examination of the curves shown in these figures indicates that the minimum variance values are clearly obtained for the condition of medium sensitivity. The same conclusion may be drawn from similar curves drawn for an aiding ratio of 3.2. Thus, it was decided to make a more intensive investigation of only the medium sensitivity case.

At this point an interesting comparison can be made between the results obtained from the case in which medium sensitivity and a 1.6 aiding ratio were used and those obtained using conventional automobile controls in a similar situation. These results, which are given in table 1, clearly show the improvement in car following obtained by the use of the tactile aiding device. For example, the average headway variance obtained from three subjects, each making three runs at 40 mph with conventional controls, was 3.5 ft^2 , while the corresponding average variance obtained using the tactile controller was 0.17 ft^2 .

TABLE 1.—AVERAGE HEADWAY VARIANCES FOR CONVENTIONAL AND "TACTILE" CONTROL

| Control type | Speed, mph | | |
|--|------------|------|------|
| | 40 | 50 | 70 |
| Conventional | 3.5 | 7.6 | 13.9 |
| Tactile controller
(Medium sensitivity; aiding ratio=1.6) | 0.17 | 0.13 | 0.18 |

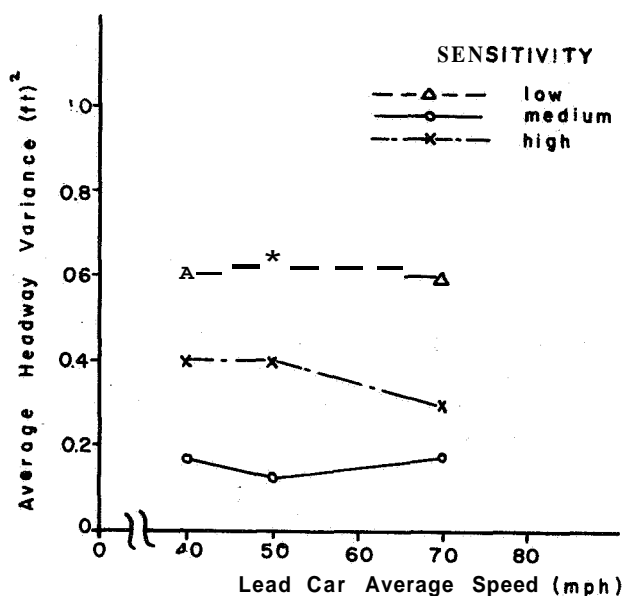


Figure 5.—Average headway variance for three values of finger sensitivity and an aiding ratio of 1.6.

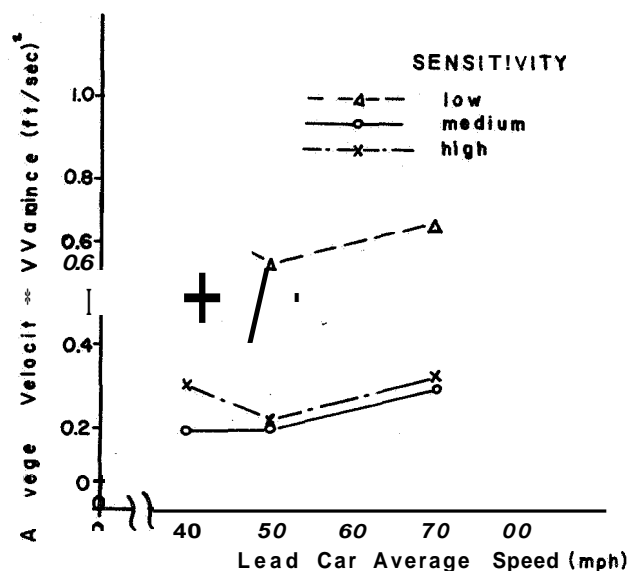


Figure 6.—Average velocity variance for three values of finger sensitivity and an aiding ratio of 1.6.

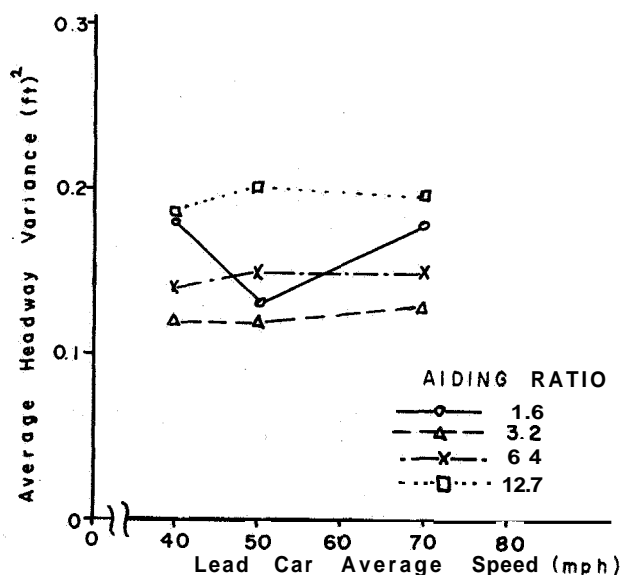


Figure 7. — Average headway variance for various aiding ratios and fixed sensitivity.

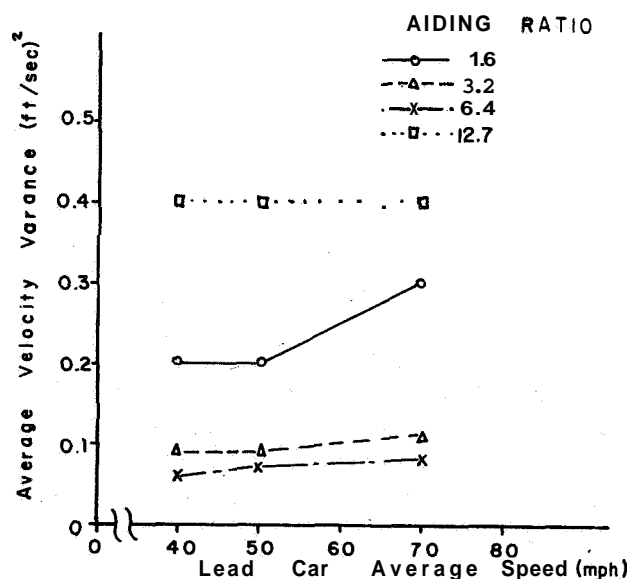


Figure 8. — Average velocity variance for various aiding ratios and fixed sensitivity,

The average headway variances obtained using medium sensitivity and four aiding ratios are shown in figure 7. The corresponding average velocity variances are plotted in figure 8. It appears from these curves that the minimum values of both headway and velocity variance are obtained for aiding ratios of 3.2 and 6.4. A later experiment in which an aiding ratio of 4.5 was used gave essentially identical results. Thus, it was concluded that no single optimum tactile aiding ratio existed and that optimum car following performance is probably obtainable for ratios in the range from 3 to 6.5. The subjects concurred in this judgment and stated that they could discern no differences when the aiding ratios were in this range. They unanimously indicated that the system was too sensitive when the aiding ratio was increased to 12.7. Two

subjects also felt that the system was too insensitive with an aiding ratio of 1.6. The subjects were unanimous in indicating that no information concerning the relative longitudinal motion of the lead car was received visually; in effect, the motions were so slight that the lead car appeared to be stationary relative to the driven car.

A MODEL OF THE DRIVER-TACTILE CONTROLLER SYSTEM

In order to ascertain if the subjects were receiving any visual cues when using the optimized tactile controller, a second experiment consisting of three parts was performed. In the first part, the procedure was identical to that previously described. The second part was identical to the first, except that the steering task was deleted. In the last part, the steering task was again deleted and the TV monitor was turned off so that the subject could not receive any visual cues. A tactile aiding ratio of 4.5 and medium finger-position sensitivity were used in each case. The averaged headway variances are plotted in figure 9 and the corresponding average velocity variances in figure 10. It appears clear from these data that little, if any, difference exists in either the headway or the velocity variances obtained from the three parts. This conclusion was substantiated by an analysis of variance which is contained in reference 10. Hence, there is strong evidence that the subjects are obtaining tactually all required information for longitudinal control of the car. If this is the case, the general system can be approximately represented by a linear transfer function relating the finger input signal and the control stick output, that is, a model of the driver-tactile controller system. Such a transfer function can be derived by use of the correlation techniques discussed by Lee (ref. 11).

In order to obtain the necessary data, each of three subjects made a 20-minute run at an average lead-car speed of 50 mph. An aiding ratio of 4.5 was used in these runs. Bode plots of the resulting linearized transfer functions are shown in figures 11 to 13. Note the similarity of these functions. In each case, a large amount of attenuation is present at very low frequencies and the slope is approximately 60 db/sec, while at intermediate frequencies the slope is zero. For $\omega > 3.5$ rad/sec, the slope is a constant 20 db/sec—a clear indication that the subjects are generating a lead term. The functions are valid linear approximations over the indicated frequency range; however, neither the very low frequency behavior nor the very high frequency behavior is correctly predicted. Specifically, one should not have complete attenuation of a dc input signal; otherwise, the operator would not respond properly to a step change in velocity of the lead car. The quantity Y_p should approach a finite non-zero value as ω approaches zero. Also, additional attenuation must be present at high frequencies, as it is not possible for a human operator to track very high frequency signals.

The general form of this transfer function for $\omega > 0.6$ is similar to those obtained from tracking studies where a visual compensatory display was used (refs. 12 and 13). Indeed, a mathematical model of the subject-tactile controller combination may be obtained which is similar to the one suggested by McRuer and Krendel (ref. 12).

A block diagram of the driver-vehicle system is shown in figure 14. The forward loop consists of the vehicle longitudinal dynamics, an aiding network, and the tactile display-operator-control stick combination. The transfer function of the aiding network is the sum of a proportional and integral term; hence, the output is a linear combination of relative velocity and headway deviation. Three open-loop transfer functions for this network—one for each of three subjects—are shown in figures 15 to 17.

Observe that the phase margin in each case is less than 60° and the gain margin is less than 6 db. Hence, while the driver-vehicle system is locally stable, it is not asymptotically stable. It is now clear that optimum tracking performance and asymptotic stability do not occur under identical conditions. Note that the overall loop gain remains constant and close to unity for $\omega > 4$. In addition, the net phase is approximately -180° ; hence the denominator of equation (2) is quite small for $\omega > 4$, and the peak value of $\left| \frac{V_2}{V_1}(j\omega) \right|$ is larger than one.

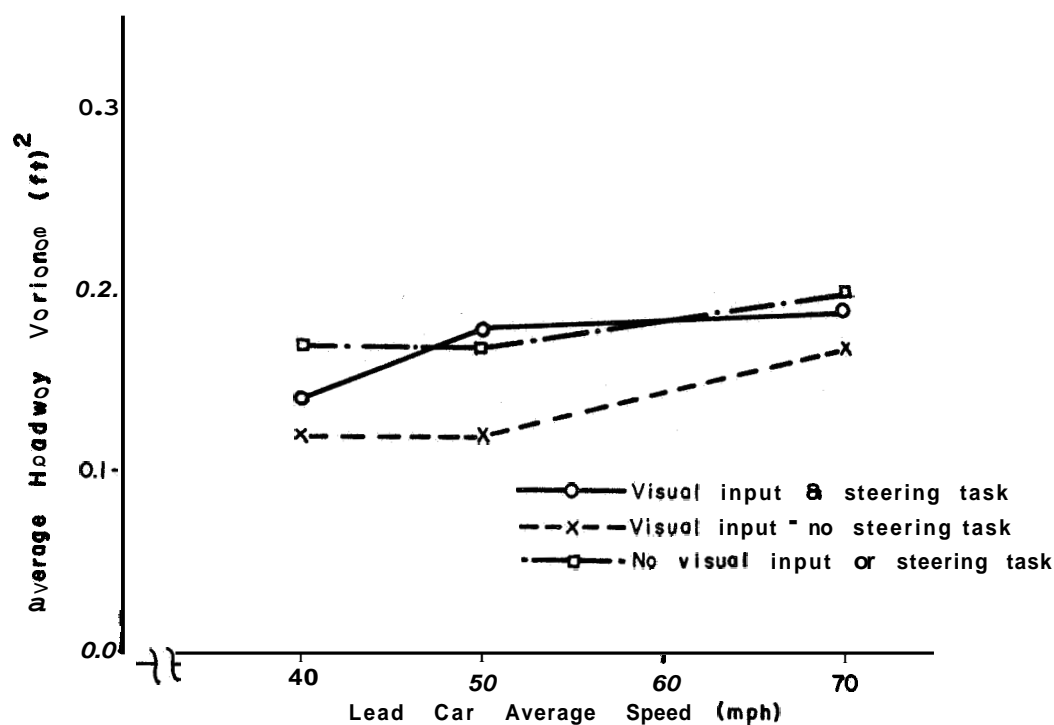


Figure 9.—Average headway variance for various driver input conditions (aiding ratio = 4.5).

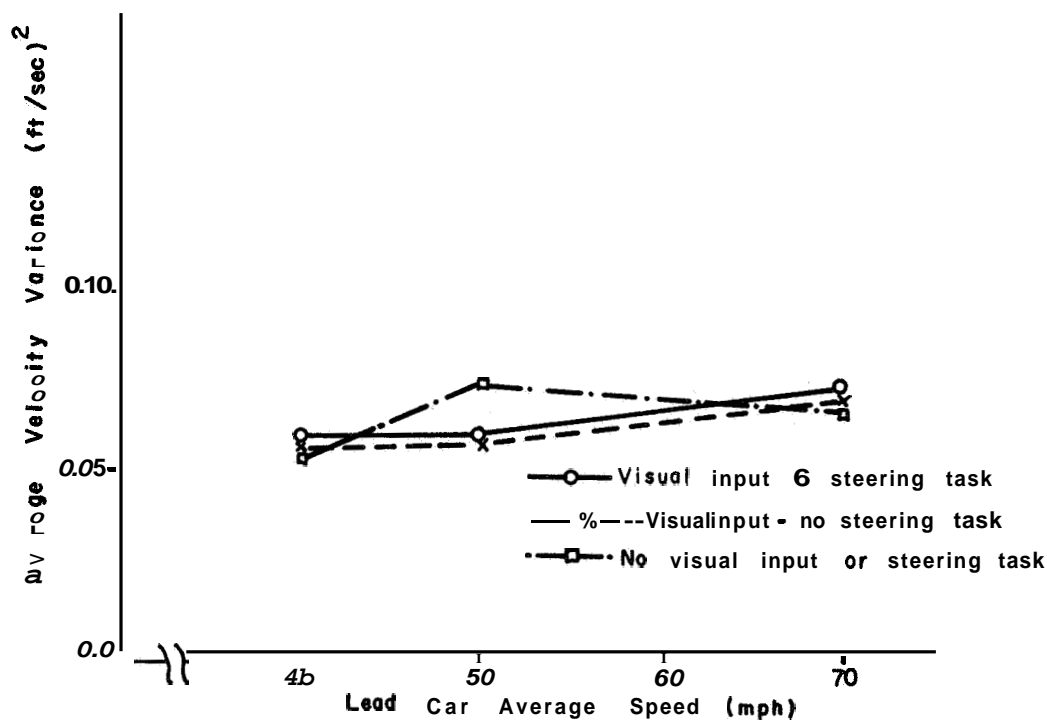


Figure 10.—Average velocity variance for various driver input conditions (aiding ratio = 4.5).

It is interesting to note that this result is partially caused by the operator lead term and effective time delay.

If equation (1) is to be satisfied, one or both of the following conditions must be met:

- (1) Lower loop gain
- (2) Additional phase lead at lower frequencies

An attempt to lower the loop gain by simply changing the gain of the aiding network was not successful, as the operator tended to increase his gain so as to maintain a constant loop gain. Other investigators have obtained a similar result from tracking studies in which a visual compensatory display was used (ref. 12).

The loop gain may be effectively decreased by the use of a deadzone in tandem with the aiding network. The use of a deadzone also simplifies the operator's task as he does not have to respond to low-level signals.

The averaged headway and velocity variances obtained from three subjects when an effective deadzone of 8 feet was used are shown in figures 18 and 19. Note the increase in headway variance as compared with the no deadzone case (fig. 7). The velocity variances are also higher than those previously obtained. However, despite the higher variances, the relative motion of the two vehicles was below the subjects' visual threshold, and the driver-tactile display combination was treated as a single-input, single-output device.

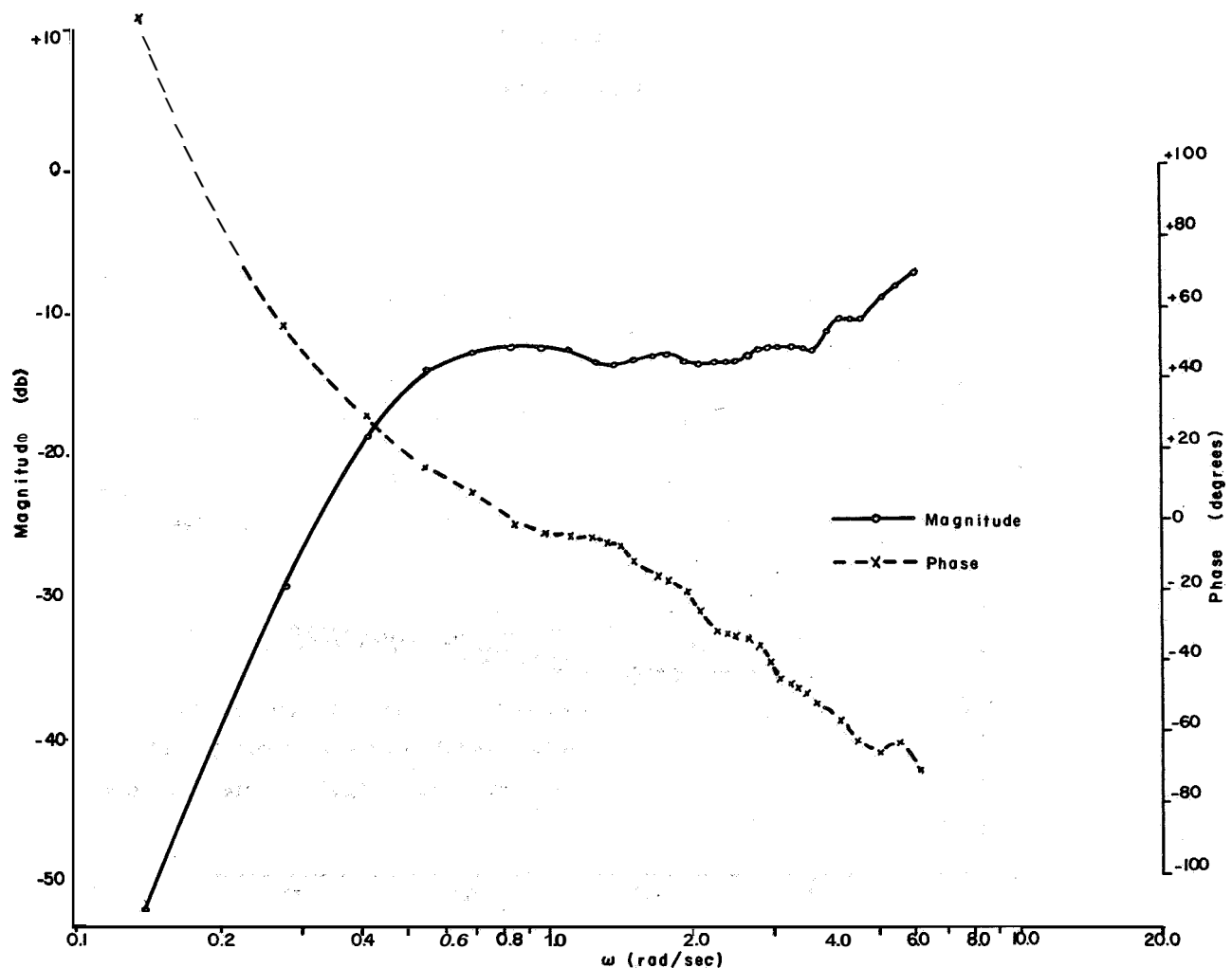


Figure 11.— Linearized transfer function of subject-tactile controller combination (subject I).

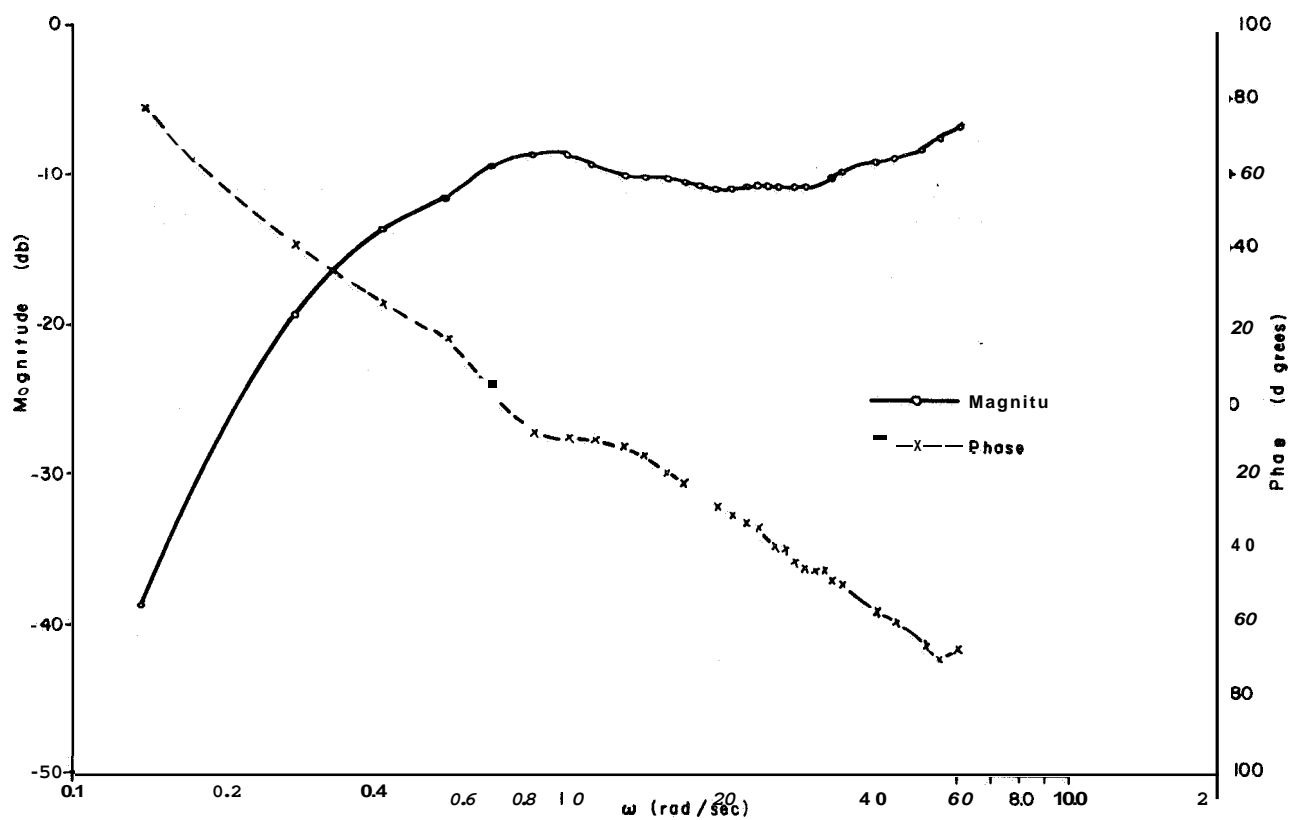


Figure 12.—Linearized transfer function of subject-tactile controller combination (subject II).

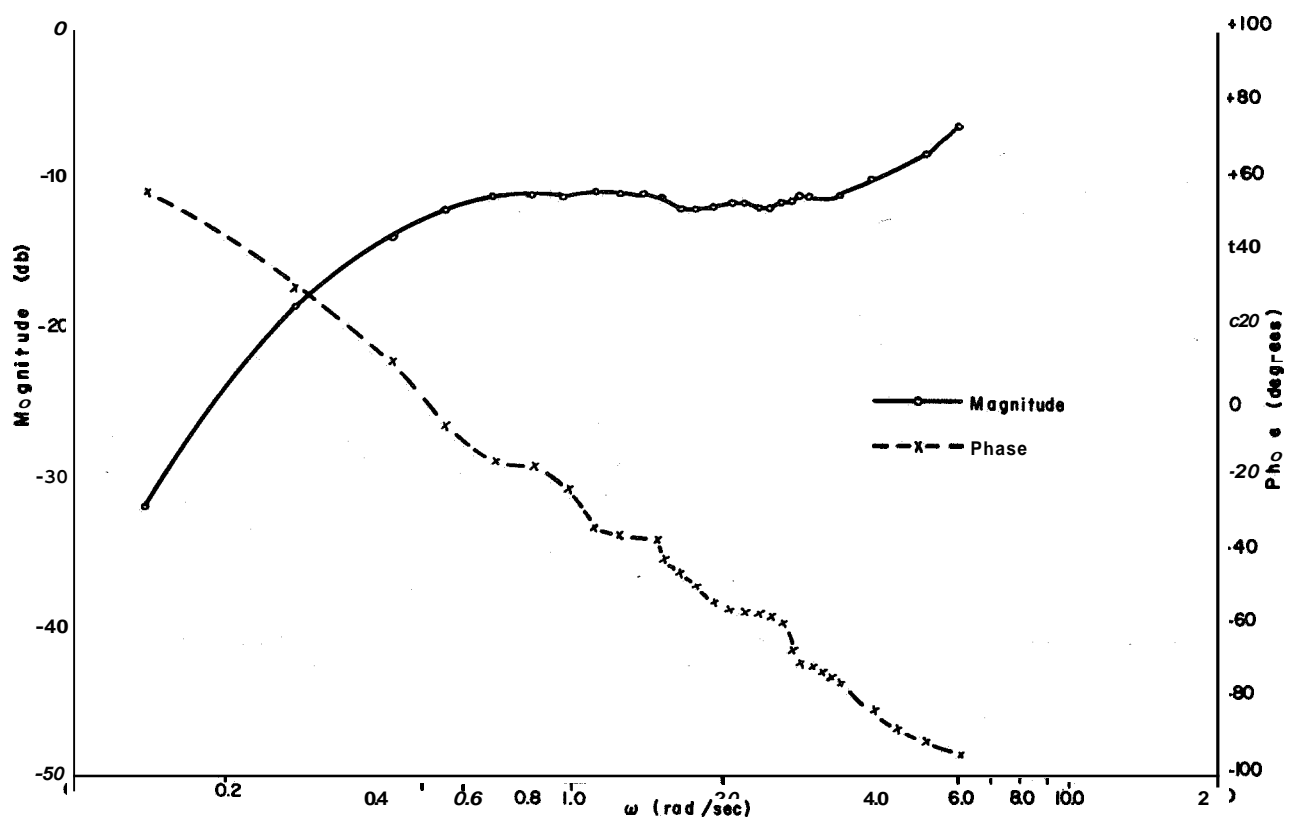


Figure 13.—Linearized transfer function of subject-tactile controller combination (subject III).

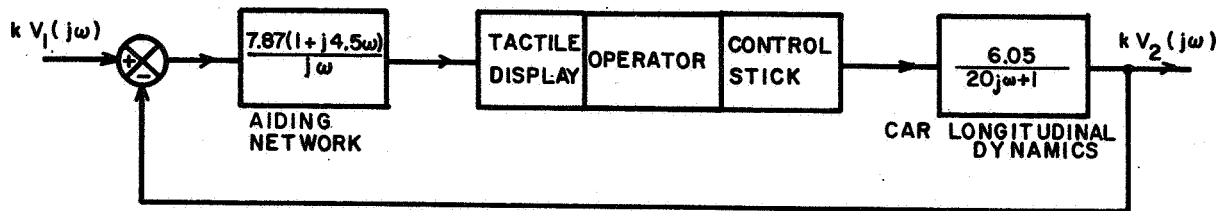


Figure 14.—Block diagram of driver-vehicle system (longitudinal mode).

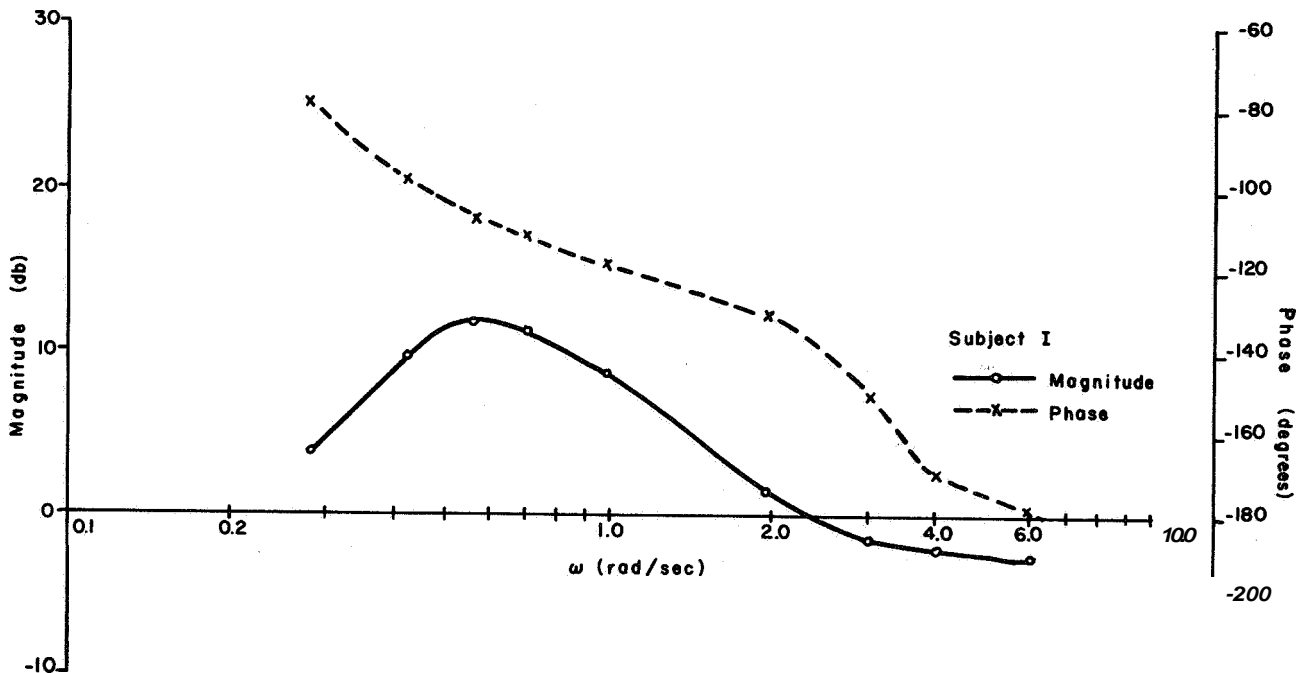


Figure 15.—Open-loop transfer function of driver-vehicle system (subject I).

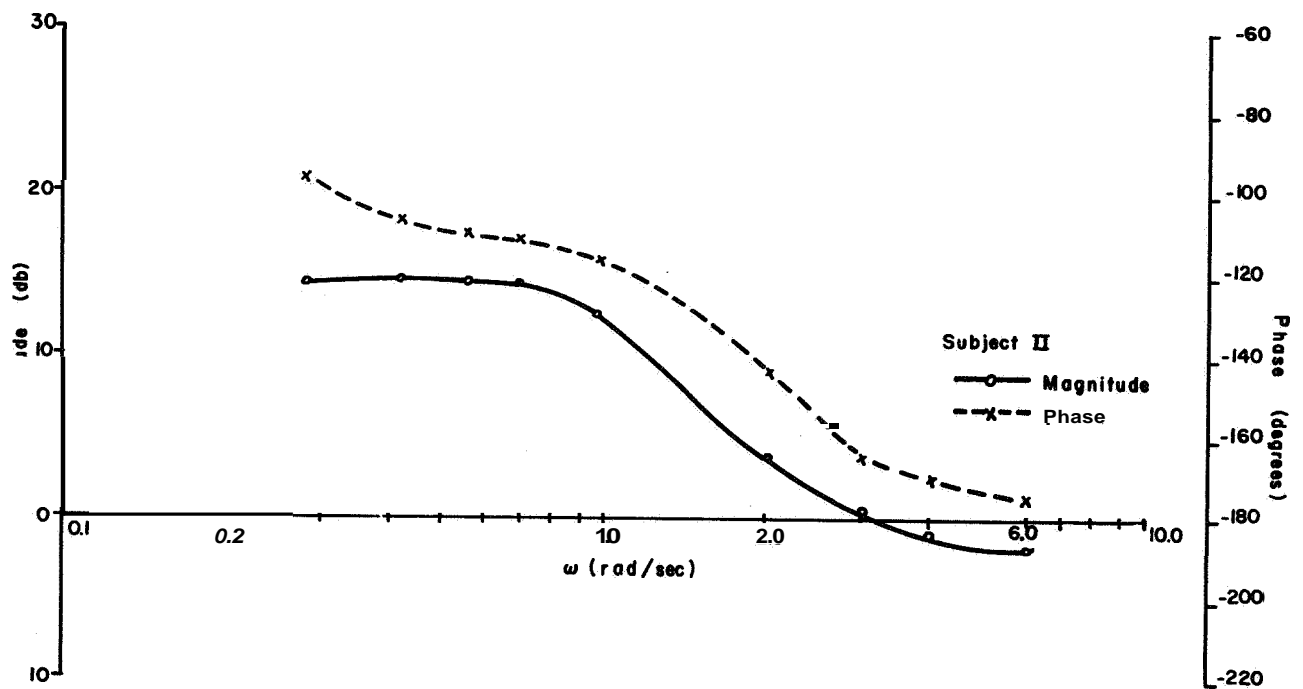


Figure 16.—Open-loop transfer function of driver-vehicle system (subject II).

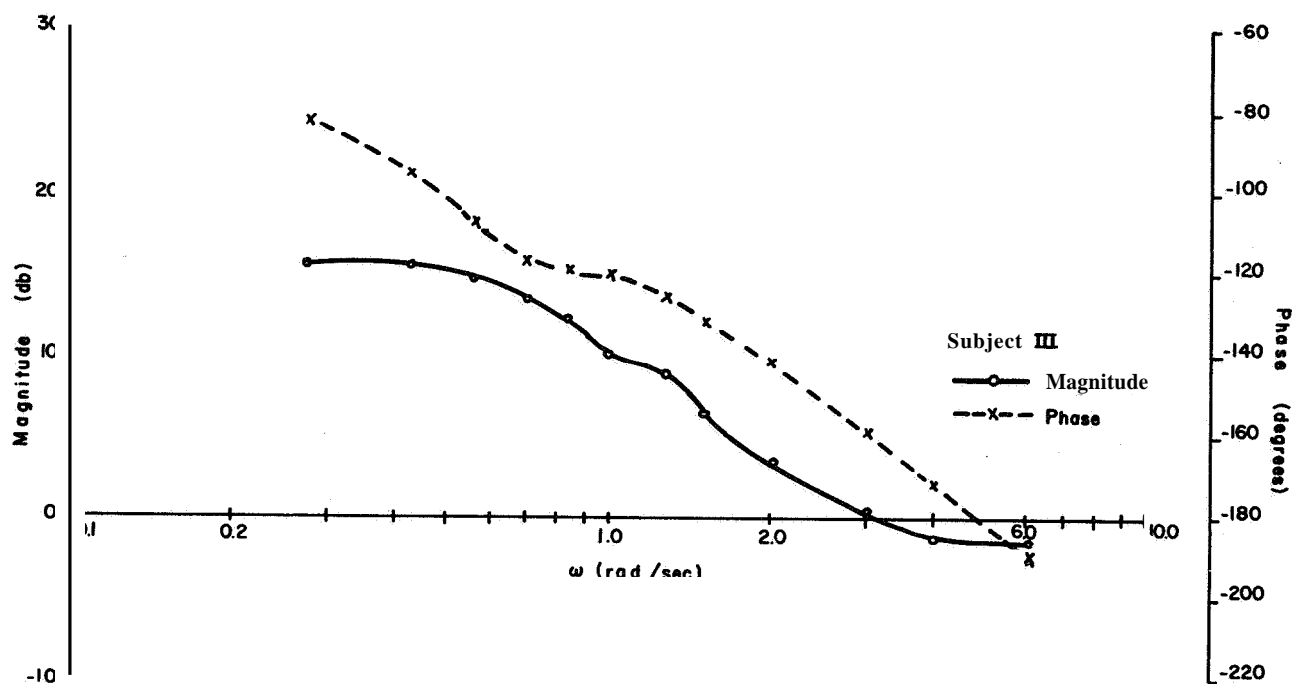


Figure 17.—Open-loop transfer function of driver-vehicle system (subject III).

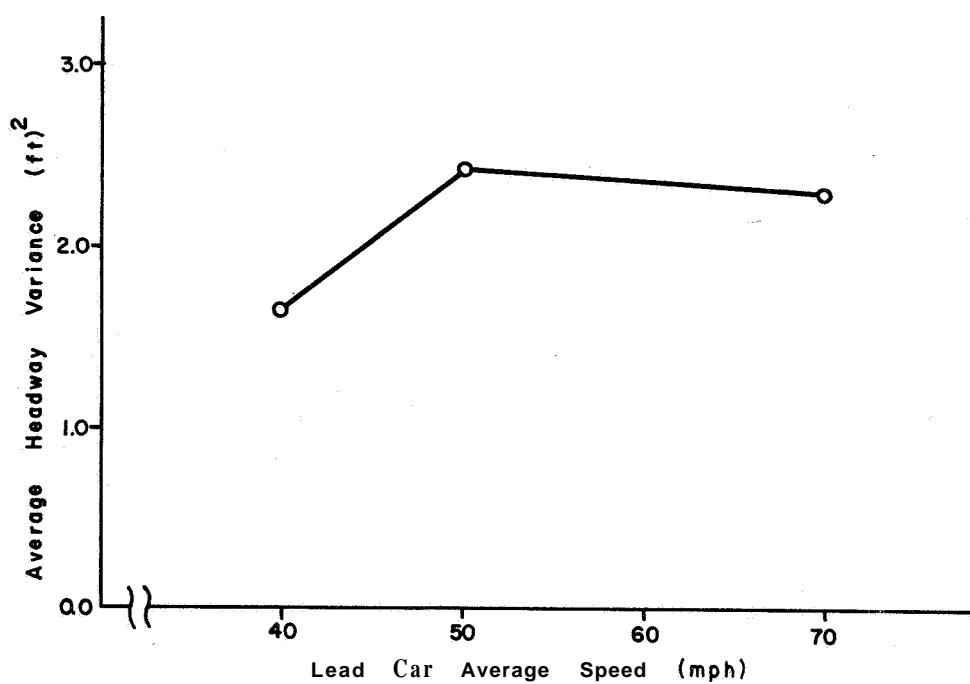


Figure 18.—Average headway variances for fixed aiding ratio and 8-foot deadzone.

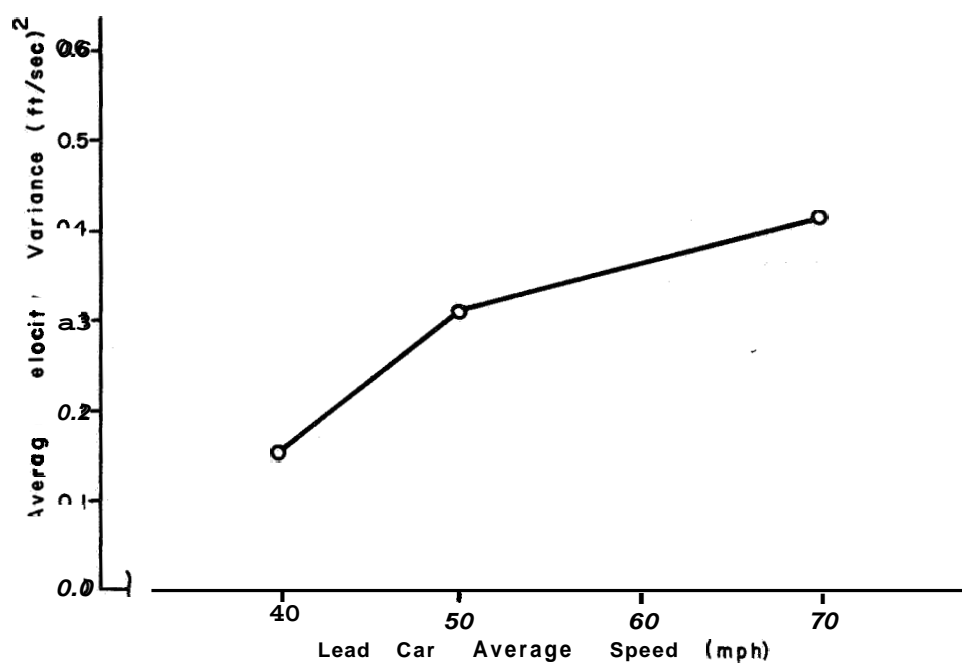


Figure 19.—Average velocity variances for fixed aiding ratio and 8-foot deadzone.

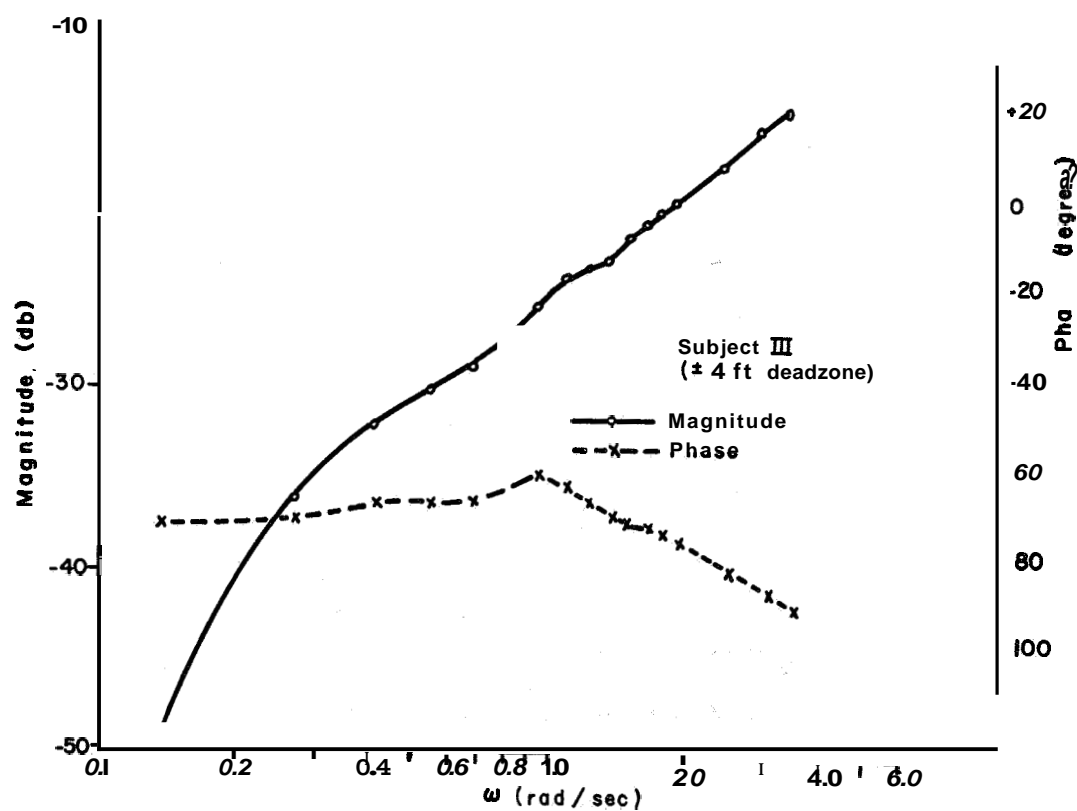


Figure 20.—Linearized transfer function of subject-tactile controller combination for deadzone case.

A linearized transfer function obtained from one subject is shown in figure 20. The shape of the magnitude curve at low frequencies is quite similar to the ones obtained under optimum tracking conditions; however, in the range above 0.6 rad/sec, the curve has a slope of 20 db/dec.

A plot of the function $\left| \frac{V_2(j\omega)}{V_1} \right|$ for this case is shown in figure 21. Note that the peak value occurs at 0.3 rad/sec and the function is less than one for $\omega > 2$. Hence, the use of the deadzone has shifted the peak to lower frequencies, and also reduced the peak magnitude. If

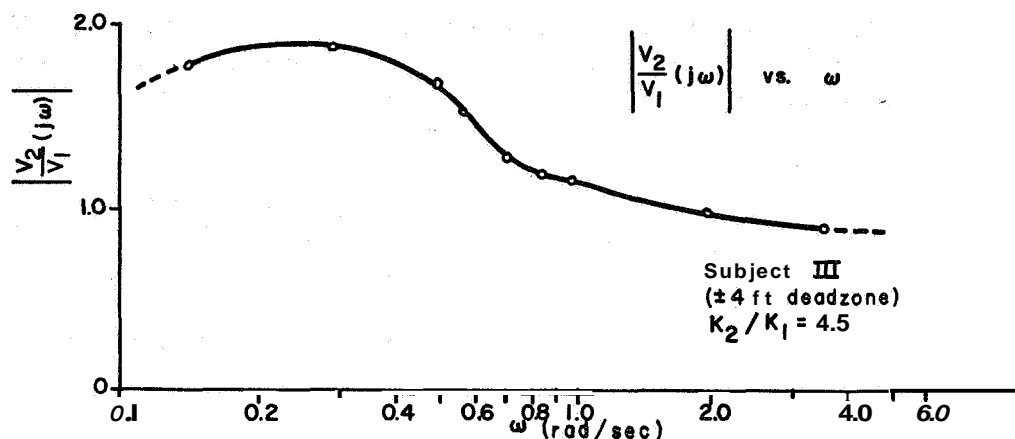


Figure 21.—Magnitude of system transfer function for deadzone case.

sufficient low-frequency phase lead and attenuation are incorporated into the system, it should be possible to reduce the peak value and satisfy the asymptotic stability criterion. Various methods for achieving this goal are currently under investigation.

CONCLUSIONS

Excellent tracking performance in simulated car following may be obtained by use of a control stick with a built-in tactile aiding device. However, tracking performance is not the only criterion of interest because it has been shown that optimum tracking and asymptotic stability do not occur under identical conditions. Indeed, the performance of the driver-vehicle system under optimum conditions leads to an asymptotically unstable system.

Asymptotic stability may probably be achieved in several ways. One approach—compensation of the driver-vehicle system—is currently being investigated by this laboratory.

REFERENCES

1. Bierley, Robert L. : Investigation of an Intervehicular Spacing Display. Highway Research Record, no. 25, 1963, pp. 58-75.
2. Montano, William B. : Tactile Aiding in a Car-Following Situation. M. Sc. Thesis, Ohio State Univ., Mar. 1967.
3. Rockwell, T. H. ; and Gantzer, D. : The Effects of Discrete Headway and Relative Velocity Information on Car Following. Paper presented at the 45th Annual Meeting of the Highway Research Board (Washington, D. C.), Jan. 17-21, 1966.

4. Fenton, Robert E. : An Improved Man-Machine Interface for the Driver-Vehicle System. IEEE Transactions on Human Factors in Electronics, vol. HFE-7, no. 4, Dec. 1966.
5. Gazis, D. C. ; Herman, R. ; and Rothery, R. W. : Nonlinear Follow-the-Leader Models of Traffic Flow. Operations Research, vol. 9, no. 4, July - August 1961, pp. 545-567.
6. Nemeth, Z. ; and Reebel, J. O. : Digital Simulation of the Car-Following Situation. Study of Electronic Devices as Traffic Aids, Transportation Engineering Center Report No. 202-1, Engineering Experiment Station, Ohio State Univ., July 1962, pp. 7-17.
7. Hanken, Albert F. : A Method and a Model for the Analysis and Description of Car-Following Performance. Systems Research Group rep. no. EES 202B-2, Engineering Experiment Station, Ohio State Univ., June 1965.
8. Barbosa, Linneu de Costa: Studies on Traffic Flow Models. Rep. no. 202A-1, The Antenna Laboratory, Ohio State Univ., Dec. 1961.
9. Todosiev, E. P. : An Automobile Simulator. AIEE paper no. DP 62-967, presented at the AIEE Empire Tri-district Meeting (Erie, Pa.), May 1962.
10. Fenton, Robert E. : A Tactile Display for Improved Car Following. Rep. EES 276A-3, Communication and Control Systems Laboratory, Ohio State Univ., Feb. 1967.
11. Lee, Y. W. : Statistical Theory of Communication. John Wiley & Sons, Inc., 1960.
12. McRuer, D. T. ; and Krendel, E. S. : Dynamic Response of Human Operators. WADC tech. rep. 56-524, Wright Air Development Center, Wright Patterson AFB, Ohio, 1957.
13. McRuer, D. T. ; and Graham, Dunstan: Pilot Models for Single- and Multi-Loop Systems with Random Forcing Functions. Rep. 134-1, Systems Technology Inc., Hawthorne, Calif., Nov. 1964.

32. Interpretation of Pilot Opinion by Application of Multiloop Models to a VTOL Flight Simulator Task

E. W. Vinje and D. P. Miller
United Aircraft Corporation

N68-15933

Analytical and flight simulator studies were conducted to develop a mathematical model for interpreting pilot opinion rating and pilot selection of optimum control sensitivity for a VTOL aircraft hovering task. Pilot longitudinal control of a VTOL aircraft hovering in turbulent air was modeled as a coupled, multiloop control task. Control characteristics of the pilot were described by adaptable gain and lead terms and by fixed lags. Pilots performed a simulated hovering task for a variety of VTOL configurations while rms hovering performance data were measured. These data were used to compute the pilot model adaptable parameters for each configuration.

Results indicate that variations in longitudinal dynamics and the intensity of pitch-attitude disturbances affected pilot opinion rating but not hovering accuracy. However, increasing the intensity of position disturbances increased hovering error while pilot opinion deteriorated. When pilot opinion deteriorated because of difficulty in controlling pitch attitude, computed pilot lead in the pitch loop increased. The pilot also disliked controlling large position disturbances, but he did not always adapt large lead terms in doing it. There was some indication that the pilot's criterion for selecting optimum control sensitivity when the longitudinal speed-stability and drag parameters were varied may have been to keep his internal pitch-loop gain constant.

Because of the emphasis placed on the development of VTOL aircraft in the last 10 years, numerous experimental VTOL handling qualities studies (refs. 1 to 5) have been conducted. Pilot opinion rating and pilot-selected (optimum) control sensitivity results from these studies have generally been presented graphically as the variation of pitch-rate damping (a stability augmentation parameter) with control sensitivity. Results from the different studies generally disagree. Tapscott (ref. 1) felt that differences in aircraft mass and moment of inertia might account for some of the disagreement. A later Princeton study (ref. 5) indicated that differences in speed-stability derivatives $M_{\dot{u}g}$ could also account for some of the discrepancies. Investigators at the United Aircraft Research Laboratories (UARL) undertook flight simulator studies to examine the effects of these and several other factors, some of which had not been given sufficient attention in the previous investigations. Results from the UARL studies (ref. 6) indicate that the stability derivatives X_u (longitudinal drag parameter) and $M_{\dot{u}g}$ the level of turbulence and the precision required in the performance of the task all have significant and systematic effects on pilot-opinion rating and optimum control sensitivity.

While the basic sources for variations in pilot rating and optimum control sensitivity appear to have been identified, insight into the process by which the pilot adapts his control characteristics to different VTOL aircraft configurations and flight conditions is lacking. An analytical description for pilot control would provide a means for defining this adaptation process. Analytical methods are also required for interpreting the pilot opinion and optimum

control sensitivity results and extrapolating to configurations or conditions not included in the previous investigations. The systematic nature of the experimental results suggest that a mathematical model can be developed **for** this purpose.

The longitudinal two-degree-of-freedom hovering task was selected as the task to be modeled. The VTOL hovering task is amenable to linear mathematical analysis and, therefore, closed-loop servo-analysis techniques can be applied to it. *Also*, the pilot's objective for the hovering task, **that is**, to minimize hovering error, corresponds to **that for** compensatory tracking tasks which have been used in the development **of** human-pilot models (ref. 7). **As a** result, some conclusions from these investigations, for example, quasilinear pilot behavior, apply to the VTOL hovering task.

Human-pilot models are fundamental to any mathematical description of the hovering task. Variations in the model parameters could provide an explanation for the observed systematic changes in pilot opinion rating and optimum control sensitivity. The describing function techniques used to develop pilot models for the one-degree-of-freedom compensatory tracking tasks, however, do not seem applicable to the two-degree-of-freedom hovering task. They do not apply because the two controlled degrees of freedom for the hovering task are coupled.

This study was undertaken to develop a model for the VTOL aircraft longitudinal hovering task, to formulate a method for computing the parameters which describe pilot control characteristics for this task from measured experimental results, and to identify any trends in computed pilot model parameters which correlate with variations in pilot rating and optimum control sensitivity.

This project is a part of the limited Aircraft Research Laboratories' corporate-sponsored research program.

SYMBOLS

| | |
|----------------|---|
| e | base for the natural system of logarithms, $e=2.71828$ |
| g | gravitational constant, 32.2 ft/sec^2 |
| I_y | moment of inertia in pitch, slug-ft ² |
| $j=\sqrt{-1}$ | |
| K | open loop gain |
| K_P | pilot adaptive gain |
| K_{P_θ} | pilot adaptive pitch-loop gain, in./rad |
| K_{P_x} | pilot adaptive position-loop gain, rad/ft |
| M | pitching moment divided by I_y , rad/sec ² |
| M_{δ_P} | longitudinal control sensitivity, (rad/sec ²)/in. |
| PM | phase margin, deg |
| PR | Cooper pilot rating (table 1) |
| q | pitch rate, rad/sec |

| | |
|----------------|---|
| s | Laplace operator, 1/sec |
| S_{u_g} | turbulence power spectrum, ft^2/sec |
| T_L | pilot adaptive rate sensing (lead) parameter, sec |
| T_N | neuromuscular lag, sec |
| T_I | pilot adaptive lag, sec |
| u | velocity perturbations along the x axis, ft/sec |
| u_g | longitudinal component of gust velocity, ft/sec |
| x | conventional longitudinal axis notation for the body-axis system |
| X | forces along x-axis, lb |
| Y_P | quasilinear pilot model |
| δ_P | longitudinal displacement of control stick, in. |
| e | pitch attitude rad |
| θ_x | position loop command angle, rad |
| σ | denotes rms value of subscripted variable for hovering task |
| σ_{u_g} | rms turbulence, ft/sec |
| τ | transport lag, sec |
| w | damped frequency of oscillation, rad/sec |
| w_B | break frequency of first-order filter for generating turbulence, rad/sec |
| ω_c | crossover frequency, rad/sec |
| ω_{ie} | effective bandwidth of turbulence, $\left[\frac{\int_0^\infty S_{u_g} d\omega}{\int_0^\infty (S_{u_g})^2 d\omega} \right]^2$, rad/sec |

Subscripts:

| | |
|--------|--|
| REF | denotes reference or base value of a variable |
| q, u | Denotes partial derivative of a parameter with respect to the subscript shown (e.g., pitch rate damping $M_q = \partial M / \partial q$) when used with M or X. Denotes pitch rate and longitudinal velocity (position rate), respectively, when used with σ . |

x, θ denotes quantities pertaining to position and pitch loops, respectively

δ_P denotes quantity pertaining to control stick activity

E denotes error between desired and actual value of a variable

$(\dot{})$ denotes first derivative with respect to time, d/dt

$(\ddot{})$ denotes second derivative with respect to time, d^2/dt^2

Limits :

db decibels, $20 \log_{10}()$ where () is an amplitude quantity, for example, Y_P

decade unit value of $\log_{10}A$, where A is a ratio of frequencies

DESCRIPTION OF THE SIMULATED VTOL HOVERING TASK

A detailed discussion of the equations of motion, the simulated turbulence, the United Aircraft V/STOL aircraft flight simulator, and methods used in performing VTOL handling qualities studies at the Research Laboratories is presented in reference 6. Consequently, only a brief summary is presented here of these topics as they apply to the precision hovering task which was modeled mathematically in this study.

EQUATIONS OF MOTION. The general form of the linearized equations of motion which describe the pitch attitude θ and longitudinal position x response of the simulated VTOL aircraft to control inputs and to turbulence are

$$\left. \begin{aligned} M_u \dot{u} + M_q \dot{\theta} - \ddot{\theta} &= -M_{\delta_P} \delta_P - M_u u_g \\ X_u u - g \theta - \dot{u} &= -X_u u_g \end{aligned} \right\} \quad (1)$$

The left side of the equations define the aircraft transient response (aircraft dynamics). The right side defines the manner in which control inputs δ_P and turbulence u_g force aircraft motion. The equations have been normalized with mass and moment of inertia. These equations were programed on the analog computer and used in conjunction with the flight simulator.

The stability derivatives M_u (generally presented as M_{u_g}), M_q , and X_u all affect the aircraft's dynamic response. The speed-stability derivative M_{u_g} is a measure of the change in pitching moment on the aircraft caused by small changes in airspeed. The longitudinal drag parameter X_u is a measure of the change in longitudinal force on the aircraft caused by small changes in airspeed. The magnitude of both of these parameters depends on the geometric, aerodynamic, and propulsive characteristics of the aircraft. When hovering, the small, rapid changes in airspeed resulting from turbulence cause pitching motion through M_{u_g} and displacement from the hovering position through X_u . Hovering position errors are also introduced by the horizontal force component which results when the pitching motion changes the orientation of the thrust vector, that is, changes g_e . Pitch rate damping M_q reduces the aircraft's pitch response to turbulence by developing a restoring moment

proportional to pitch rate. The longitudinal control sensitivity $M_{\delta P}$ is the initial pitch acceleration per inch of control stick displacement.

SIMULATION OF TURBULENCE. The simulated turbulence was generated by passing the output of a low-frequency random-noise generator through a first-order filter. The filter had a break-point frequency, ω_B , of 0.314 rad/sec. That is, the signal was unattenuated out to ω_B , then attenuated at 20 db/dec for higher frequencies. The resulting simulated turbulence had a power spectral density given by

$$S_{u_g}(\omega) = \frac{2\omega_B \sigma_{u_g}^2}{\omega_B^2 + \omega^2} \quad (2)$$

where σ_{u_g} is the root-mean-square value of the random component of the wind. A value of $\sigma_{u_g} = 5.14$ ft/sec was used for this study. The simulated turbulence is the same as that simulated in the flight tests of a variable-stability tandem-rotor helicopter reported in reference 5. In that study the pilots described this simulated turbulence as corresponding to a medium-turbulent day.

V/STOL AIRCRAFT FLIGHT SIMULATOR.—The hovering task was performed in the V/STOL aircraft flight simulator located in the Analog Computation Laboratory of the United Aircraft Research Laboratories. This facility consists of a full-scale, fixed-base, Sikorsky S-61 helicopter cockpit with a Norden Contact Analog display system. Figure 1 shows the



Figure 1.—V/STOL aircraft flight simulator.

interior of the cockpit, the conventional helicopter-type controls, and the contact analog display. Aircraft pitch attitude was controlled with a conventional helicopter control stick having inertia but no force gradient. Longitudinal position was controlled by changing pitch which changed the orientation of the thrust vector and produced longitudinal forces.

The Norden Contact Analog system provided the pilot with an inside-out video picture of the motion of the aircraft relative to the earth and sky. Figure 2 shows the contact analog display for the VTOL hovering task. The cross indicated the position of the nose of the aircraft relative to the ground and hence remained fixed at the center of the viewing screen. The square traversed vertically on the screen and was used as a sensitive longitudinal position indicator. When the square was at the center of the screen (coincident with the cross), the aircraft was located directly

over a reference hovering position such that the ground position indicator (also shown in fig. 2) was approximately 135 feet ahead of the aircraft for the simulated hovering altitude of 40 feet.

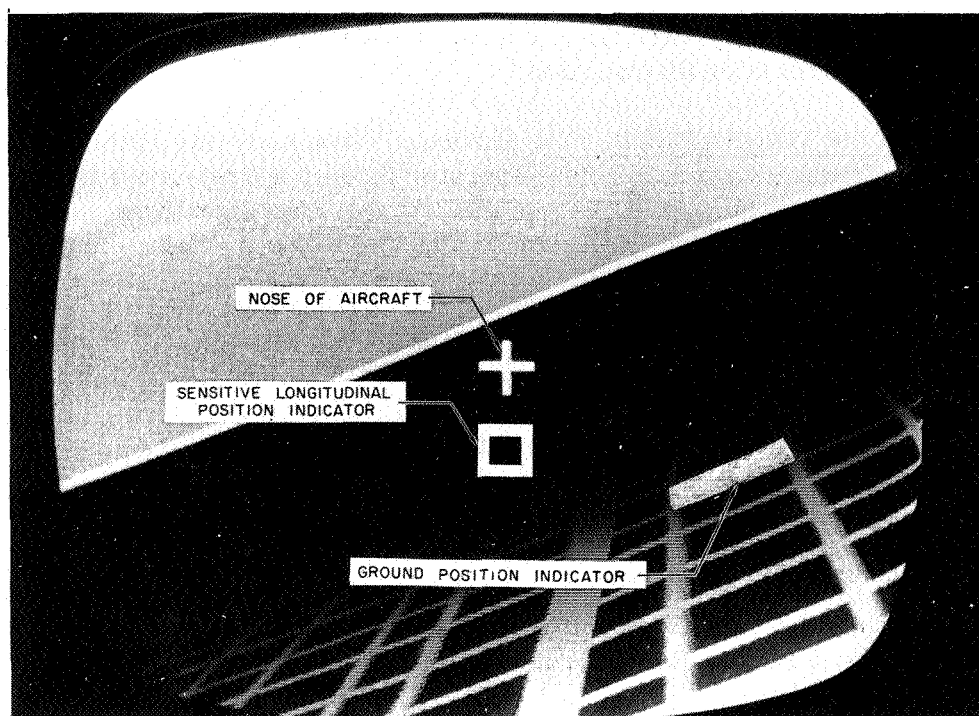


Figure 2.—Contact analog display for hovering task.

As the aircraft maneuvered 50 feet forward and rearward, with respect to the nominal reference hovering position, the square traveled to the bottom and top edges of the screen, respectively. Figure 2 shows the aircraft rolled to the right. The roll and lateral position degrees of freedom were included in the task simulation. However, the roll motion was well damped, lateral disturbances due to turbulence were small, and little effort was required to control roll attitude and lateral position. As a result, the presence of the lateral degrees of freedom did not appreciably affect the pilot's performance or opinion of the longitudinal hovering task.

INSTRUCTIONS TO PILOTS. The pilots performed the hovering task for 12 different VTOL aircraft configurations. The ranges of stability derivatives investigated were :

$$\left. \begin{aligned} -5.0 &\leq M_q \leq -1.0 \\ -0.3 &\leq X_u \leq 0 \\ 0 &\leq M_{u_g} \leq 1.0 \end{aligned} \right\} \quad (3)$$

These derivatives were varied about a "common point" ($M_q = -3.0$, $X_u = -0.1$, $M_{u_g} = 0.667$). For each configuration the pilot was instructed to select the value of control sensitivity M_{δ_P} that he considered optimum for the hovering task. The control sensitivity could be adjusted using a lever in the cockpit. He was then instructed to attempt to minimize longitudinal hovering position error, that is, attempt to keep the cross centered in the square, and assign a pilot rating based on the difficulty in hovering accurately. The pilot was also required to

to hover while attempting to minimize hovering position error for 10 runs of 100 seconds duration each. For each of these runs, rms longitudinal hovering position error σ_x , rms longitudinal velocity σ_u , rms pitch attitude σ_θ , and rms pitch rate σ_q were measured. As discussed later, these rms values were subsequently used in computing the adaptive pilot model parameters. The rms results for all values of a particular stability derivative were generally measured on the same day.

DESCRIPTION OF THE HOVERING TASK MODEL

This section of the report contains a description of the model which was developed for the longitudinal hovering task. The mathematical representation for the human pilot, incorporation of the pilot model into an overall task model, and an example of the model's operation in turbulent air are discussed.

HUMAN-PILOT MODEL. The human-pilot model was of primary importance in this study. It was felt that variations in the parameters used to describe the pilot would provide insight into the process by which the pilot adapts to changes in flight conditions. It was also thought possible that these parameter variations might indicate the reasons for trends in pilot opinion rating and optimum control sensitivity. Consequently, an analytical representation which was reasonably accurate and complete was desired for the pilot. However, it was also felt that this representation should be simple enough to permit the development of task model equations which did not include complex nonlinear terms.

One-degree-of-freedom compensatory tracking tasks have been used frequently in the study of human-pilot dynamic behavior (refs. 7 and 8). The fundamental characteristic of the longitudinal hovering task is very similar to that of one-degree-of-freedom compensatory tracking tasks. That is, the human pilot is continuously striving to minimize the error, caused by some disturbing function, between the actual location of a controlled degree of freedom and some fixed reference location. For the longitudinal hovering task the pilot is controlling two degrees of freedom, and they are coupled. His primary concern is generally with position, that is, he desired to minimize the error between his actual hovering position and the desired or reference position. However, in the process of minimizing hovering position error the pilot must also suppress disturbances in the other (pitch) degree of freedom. When pitch attitude is other than its reference value of zero, hovering position errors are introduced because the thrust vector is no longer normal to the Earth and a component of longitudinal force results (a coupling effect). The longitudinal hovering task can thus be considered to be similar to two coupled compensatory tracking tasks.

Because of the similarity between the hovering task and the compensatory tracking task, it appeared that the quasilinear model developed to describe pilot control of the latter task would also apply to hover control. It was shown that, for one-degree-of-freedom compensatory tracking tasks (refs. 7 and 8), this model adequately describes the fundamental human control characteristics for a broad range of controlled element dynamics, disturbing-function amplitudes, and disturbing-function frequency content. In reference 9 the same conclusions regarding the adequacy of the quasilinear model were arrived at for an uncoupled two-degree-of-freedom compensatory tracking task. The quasilinear pilot model also meets the requirements for simplicity which were outlined. Therefore, it was considered a logical choice for this initial study.

The form of the quasilinear pilot model used in the study is given in Laplace transform notation by

$$Y_P = \frac{K_P (T_L s + 1) e^{-\tau s}}{(T_N s + 1)} \quad (4)$$

Pilot gain, K_p , defines the pilot's control response to an error in the magnitude of a controlled variable. The lead term $(T_L s + 1)$ is an indication of the pilot's control response to the rate of change in error of the controlled variable. The parameters K_p and T_L are adaptive. When controlled element dynamics and/or disturbance function characteristics change, the pilot adapts his gain and lead to the new conditions in an attempt to maintain stable control and achieve desirable system-response characteristics. There is some evidence (ref. 7) that the pilot can also adapt a first-order lag term, $(T_I s + 1)^{-1}$, if this is necessary to achieve the desired system response. For the controlled element (VTOL aircraft) dynamics and disturbance function (turbulence) characteristics considered in this study, the adaptation of lag by the pilot would not appear to improve hovering accuracy. It was assumed, therefore, that the pilot would not use lag in his control of the simulated task and this term was not included in the model.

The model factors $(T_N s + 1)^{-1}$ and $e^{-\tau s}$ define inherent pilot lags. The transport lag or pure time delay described by the $e^{-\tau s}$ term represents an accumulation of delays encountered in the transmission of information from the eye to the brain, decision making, and transmission of information from the brain to the spinal cord and finally to the muscles. The neuromuscular lag represented by the approximation $(T_N s + 1)^{-1}$ describes the lag involved in moving the wrist and arm to a commanded position after the command signal arrives at the muscles. Results in reference 7 indicate that τ can be assumed to be constant for a given individual, although it may vary among individuals. A representative value for τ is 0.09 second. It is presently believed that the frequency content of the disturbing function has the most significant effect on T_N . A monotonic decrease in T_N with increasing disturbing function bandwidth was observed in reference 7.

HOVERING TASK MODEL INCORPORATING THE PILOT MODEL. Following selection of the human-pilot model, efforts were directed towards developing a model for the simulated hovering task which incorporated the quasilinear pilot model. The compensatory tracking tasks discussed previously have been described analytically (refs. 7 to 9) as closed-loop servo-control systems. This is reasonable since, for these tasks, the pilot generates control commands based on the error between the actual and reference location of the controlled degree-of-freedom. That is, the pilot's control action is dependent on the output of the controlled element. Because the longitudinal hovering task appears to be a combination of one-degree-of-freedom compensatory tracking tasks, it seemed that the hovering task could be described as a multiple closed-loop servo-control system. In reference 10, Stapleford represented the longitudinal hovering task in this manner, calling his description of the task a "parallel loop" model. In later discussions with Stapleford he indicated that a better representation for the task might be a "series loop" model. In his series loop representation it is assumed that the pilot, operating only on the magnitude of position error (not the rate of change of this error), mentally generates a pitch-attitude angle which is necessary to develop the longitudinal force needed to reduce the position error. The pilot mentally combines this required pitch angle with the present pitch-attitude angle of the vehicle and generates a control input to null the error between the two. It was felt at the Research Laboratories that the series-loop model might be basically adequate, but that the pilot's formulation of the attitude angles required to control position was also dependent on position rate. After lengthy observation in the flight simulator of pilots' control response and discussions between the pilots, the hovering task model shown in figure 3 was developed.

A hypothetical hovering situation will illustrate the operation of the task model. Assume that the aircraft is hovering directly over x_{REF} with zero pitch attitude angle and that turbulence (positive u_g) encounters the aircraft. The aircraft will begin to pitch up and translate rearward because of the action of turbulence on $M_u g$ and X_u , respectively. Because the pitch

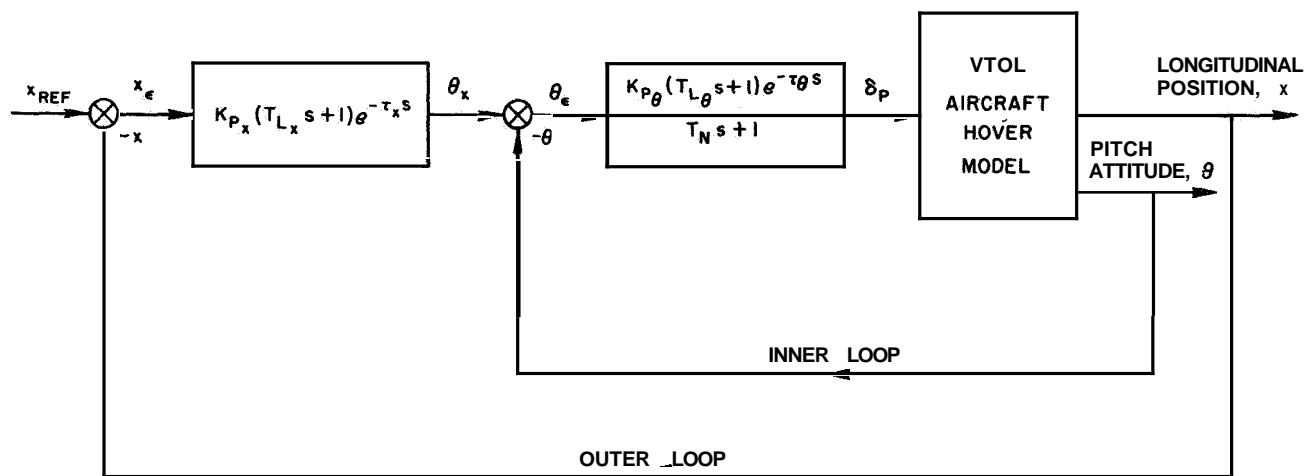


Figure 3.—Series loop model for pilot longitudinal control in hover.

response is higher frequency (larger bandwidth) than position response, the change in pitch will be observed before a hovering position error is noticeable (the pilot's ability to sense position motion is dependent on altitude). The command pitch angle θ_x will be initially zero; the pitch-loop model will operate to null the pitch angle error and error rate and, therefore, generate forward stick motion δ_P . The output δ_P will lag the pitch error input θ_e because of the information processing delay $e^{-\tau_\theta s}$ and also the neuromuscular delay $(T_N s + 1)^{-1}$. The stick motion δ_P will initiate a downward change in aircraft pitch attitude. By this time turbulence, acting through X_u , will noticeably have begun to blow the aircraft back from the desired hovering position, x_{REF} . The x or outer loop pilot model will operate on the position error and error rate to develop a θ_x command. The commanded θ_x lags x_e because of the information processing involved. No neuromuscular lag is included in the x loop pilot model as no wrist or arm motion is necessary in the development of θ_x . The new θ_x is combined with the existing pitch angle to develop a new pitch error θ_e . The pitch-loop pilot model then operates again on θ_e and $\dot{\theta}_e$.

By routing the position information in sequence through both pilot models, a component of δ_P (neglecting lags) equal to $(K_{P_x} T_{L_x} K_{P_\theta} T_{L_\theta}) s^2 x$ is developed. This indicates that stick

inputs are generated in response to position acceleration ($s^2 x = \ddot{x}$). Effectively, this term describes the information coupling between the pitch and position loops. It has been pointed out that the bandwidth of the pitch loop is larger than that of the position loop. The pitch loop, therefore, responds to higher frequency inputs. When turbulence strikes the aircraft, pitch attitude responds more rapidly than aircraft position. An experienced pilot observes the pitch response to gusts and realizes that he will be blown off in position unless he acts quickly to counteract the turbulence. Pilots flying the simulated task made stick inputs to counteract expected position errors before position rates could be discerned. Through this information coupling they were effectively operating on position acceleration. Their opinion of the hovering task deteriorated when the coupling was removed for configurations having high pitch-rate damping and/or low-speed stability.

COMPUTATION OF THE PILOT MODEL ADAPTIVE PARAMETERS.—The approach commonly used to measure pilot-describing functions (quasi-linear models) for one-degree-of-freedom compensatory tracking tasks involves the computation of a cross-spectral density function between the disturbance function and the pilot's control-stick motion. It does not appear that the spectral analysis approach would yield unique pilot model describing functions for each of the loops controlled in the coupled hovering task. Low-frequency turbulence introduces errors in both pitch and position which the pilot is forced to respond to using a single control. Spectral analysis would not determine which portion of the pilot's control activity was in response to an error in pitch and which was in response to position. Therefore, an alternate method had to be developed to compute the pilot model parameters for the hovering task model.

It was believed that the statistical (rms) hovering performance data from the piloted flight simulator could be used to compute the adaptive parameters. The hovering task model was linearized by using a first-order Pad6 approximation for the pilot model transport lags. The Pad6 approximation is given by

$$e^{-\tau s} \approx \frac{-\tau s/2 + 1}{7s/2 + 1} \quad (5)$$

Transfer functions were then written for the hovering task model which defines the closed-loop pitch, pitch rate, position, and position rate response to turbulence (θ/u_g , q/u_g , x/u_g , and u/u_g , respectively). For these computations it was assumed that the pilot closed the pitch loop before closing the position loop. The equations describing the rms response of the task model to turbulence, that is, analytical expressions for σ_θ , σ_q , σ_x and σ_u , were developed using these closed-loop transfer functions and the analytical representation for the turbulence. For example, it can be shown that

$$\sigma_\theta^2 = \frac{2\sigma_{u_g}^2}{\omega_B} I_{\theta\theta} \quad (6)$$

where

$$I_{\theta\theta} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \left| \frac{\theta}{u_g}(s) \cdot \left(1 + s/\omega_B\right) \right|^2 ds \quad (7)$$

The integration indicated in equation (7) can be accomplished with the tables contained in appendix E of reference 11. Equation (6) can also be used to describe σ_q , σ_x , and σ_u if the appropriate closed-loop transfer function (q/u_g , x/u_g , and u/u_g , respectively) is substituted into equation (7). For a given VTOL aircraft configuration, all task model parameters except the pilot model adaptive parameters were assumed to be fixed. Thus four rms equations were provided in terms of the four pilot model adaptive parameters which were to be computed, and a unique determination of the parameters was insured.

A digital computer iteration program which included the four equations describing σ_θ , σ_q , σ_x , and σ_u was written. The measured values of σ_θ , σ_q , σ_x , and σ_u from the flight simulator were input to the program, and the four pilot model adaptive parameters were adjusted until the computed rms values matched the measured values to within 0.5 percent. The measured rms performance data were the average of 10 runs of 100-second duration each for each VTOL aircraft configuration flown by the simulator pilots.

As indicated above, the pilot model adaptive parameters were considered to be the only task variables when the pilot was flying a given configuration. That is, the pilot's transport

lag T and neuromuscular lag T_N were considered constant for the simulated task. As discussed previously, this assumption definitely holds for T , but T_N has been observed to vary with changes in disturbing function frequency. The turbulence characteristics were held constant throughout the study, but the filtering effect of the aircraft pitch dynamics did alter the turbulence bandwidth evident to the pilot. However, the results in reference 7 indicate that the variations in T_N can be considered second-order effects relative to the expected changes in pilot adaptive parameters. The value of T_N used (0.35 sec) was that given in the results of reference 7 for $\omega_{ie}=1$ rad/sec where the pilot was controlling dynamics (K/s^2 dynamics) which required him to supply lead. These conditions most closely parallel the pitch dynamics and turbulence characteristics (prior to the aircraft filtering effects) used in this study.

By using rms pitch rate instead of rms control activity as one of the measured functions used in computing the pilot model adaptive parameters, the effects of pilot higher frequency nonlinearities on the computed parameters was reduced. These pilot higher frequency nonlinearities were filtered by the aircraft dynamics and therefore did not have as significant an effect on measured σ_q as they did on measured σ_{δ_P} . These higher frequency nonlinear effects were successively less significant on σ_θ , σ_u , and σ_x since the aircraft filtering effects increased with each integration between control input and eventual controlled variable response.

VALIDATION OF PILOT MODEL PARAMETER COMPUTATION TECHNIQUES.—A simulation was developed to validate the technique for computing the pilot model adaptive parameters from measured rms values of aircraft closed-loop response to turbulence. The hovering task model was programed on an analog computer with a preselected set of adaptive parameters in the pilot models. The analog task model was subjected to the same simulated turbulence used to excite the aircraft configuration flown on the piloted simulator. Values of σ_q , σ_θ , σ_u , and σ_x were measured and the average of these functions for 10 and also 20 runs of 100-second duration each were computed. The technique described previously for computing the adaptive parameters from the measured rms results was applied. The adaptive parameters programed in the analog computer simulation and those recovered from 10- and 20-run averages of rms results are shown in the following table:

| Adaptive parameters | Programed parameters | Recovered parameters | |
|---------------------|----------------------|----------------------|---------|
| | | 10 runs | 20 runs |
| K_{P_θ} | 40.2 | 30.7 | 35.3 |
| K_{P_x} | - .0272 | - .0268 | - .0275 |
| T | .170 | .188 | .182 |
| T_{L_x} | .401 | .452 | .395 |

It is felt that these results validate the techniques used to recover the adaptive parameters. The results show that the pilot adaptive parameters would have been defined somewhat more precisely had an average value from 20 runs been used instead of 10 runs used for the results presented in this study. However, better definition of the parameters was not necessary to identify trends in the pilot model parameters with pilot opinion.

RESULTS

(1) Piloted flight simulator rms hovering performance data for the longitudinal two-degree-of-freedom hovering task were duplicated analytically by modeling the task as a multi-loop configuration which included adaptive quasilinear pilot models. By duplicating the measured rms data it was possible to compute the pilot model adaptive parameters for this coupled control task.

(2) The pilot adapted to variations in aircraft dynamics and intensity of pitch-attitude disturbances with no change in hovering accuracy even though pilot opinion rating was affected by these variations.

(3) The pilot was not able to adapt to increases in the intensity of turbulence-induced hovering position disturbances without experiencing a nearly proportional increase in hovering error. Pilot opinion also deteriorated as the position disturbances were increased.

(4) Because of the larger bandwidth of the pitch-attitude loop, the pilot was able to predict position disturbances by observing pitch response to turbulence. The pilot generated control stick inputs to attenuate expected rates of change in hovering position before these position rates could be observed. When the information coupling between pitch and position response could not be observed because of low-speed stability $M_{u\dot{g}}$ or high pitch rate damping $M_{\dot{q}}$, pilot opinion deteriorated.

(5) Results show some correlation between pilot opinion rating and computed pilot model lead. When pilot opinion deteriorated because of difficulty in controlling pitch attitude, computed pitch-loop lead increased. However, when opinion deteriorated because of difficulty in controlling position disturbances, computed position-loop lead did not always increase. For example, pilot opinion deteriorated as X_u became larger, even though position-loop lead decreased.

(6) The pilot was not willing to make the large, rapid changes in pitch attitude which were necessary to supply appreciable position-loop lead for configurations having large turbulence-induced position rates. Results indicate that the pilot preferred to relax his position-loop control as position disturbances increased and thereby reduce the trend towards violent attitude motion.

(7) When the aircraft speed-stability parameter $M_{u\dot{g}}$ and drag parameter X_u were varied, the computed pilot model pitch-loop gain, K_{P_θ} , remained nearly constant. For these variations in $M_{u\dot{g}}$ and X_u , the pitch-attitude oscillatory dynamics (complex roots of characteristic equation) changed, but the first-order root remained essentially fixed. It is possible that when the pitch dynamics exhibit these characteristics the pilot's criterion for selecting optimum control sensitivity is to maintain K_{P_θ} constant.

(8) Previously published results for one-degree-of-freedom compensatory tracking tasks indicate that the human operator frequently adapts to the controlled element dynamics so as to produce a -20 db/dec slope of the magnitude Bode plot in the region of crossover frequency. Results for the studies reported in this paper show position-loop slopes which were approximately -30 db/dec for most controlled configurations. The slope of the pitch-loop magnitude Bode plot at crossover frequency was consistently higher than that for the position loop. Pitch-loop slopes ranged from -30 db/dec for easily controlled configurations to -38 db/dec for configurations which were difficult to control. It appears that the position-loop slopes for the present study were more negative than -20 db/dec because the controlled element dynamics were higher order than those of the previous studies.

DISCUSSION OF RESULTS

Results are presented for variations in the simulated aircraft pitch-rate damping, speed-stability, and drag parameters. For each parameter, rms hovering performance data and pilot ratings (table 1) obtained in flight simulator experiments are shown. The adaptable pilot-model parameters computed from the performance data are presented and possible correlations with pilot opinion ratings and performance are discussed. In general, results are presented for only one pilot, although a comparison with the performance data and recovered pilot model parameters of a second pilot is shown for one configuration.

RESULTS FOR VARIATIONS IN PITCH RATE DAMPING, M_q

Flight Simulator Data.—Optimum control sensitivities, pilot ratings, and rms hovering performance data were obtained for three values of pitch damping, M_q (-1.0, -3.0, and -5.0). The remaining aircraft stability derivatives and turbulence characteristics were held constant at their "common point" values when M_q was varied ($M_{u\dot{g}}=0.667$, $X_u=-0.1$, $\sigma_{u\dot{g}}=5.14$, and $\omega_B=0.314$). The open-loop aircraft transfer functions for these configurations are shown in table 2. The pilot selected the following values of optimum control sensitivity:

| M_q | M_{δ_P} |
|-------|----------------|
| -1.0 | 0.369 |
| -3.0 | .431 |
| -5.0 | .493 |

The pilot ratings and rms hovering data for these configurations are presented in figure 4.

The effect of M_q on the measured rms values of pitch attitude, pitch rate, position, and position rate (σ_θ , σ_q , σ_x , and $\sigma_{\dot{u}}$) is shown in figure 4. As pitch rate damping increased, σ_q decreased while σ_θ , σ_x , and $\sigma_{\dot{u}}$ remained nearly constant. Control stick activity, σ_{δ_P} , also decreased as M_q was increased. Cooper pilot rating (PR) was unsatisfactory at M_q of -1.0, became satisfactory at -3.0 and deteriorated slightly at -5.0. The improvement in rating when M_q was increased to -3.0 was accompanied by a decrease in pilot control stick activity. As M_q was increased to -5.0 the trend in pilot rating reversed and it deteriorated.

Computed Pilot Model Parameters.—The pilot model adaptive parameters computed from these rms results are plotted in figure 5. A large decrease in computed pitch-loop lead, T_{L_θ} , is noted as M_q was increased from -1.0 to -3.0. A significant improvement in PR also occurred. With M_q of -1.0, pitch attitude is difficult to control and very responsive to turbulence. Pitch response to control inputs is more predictable and disturbances introduced by turbulence are reduced at M_q of -3.0. At $M_q=-5.0$, T_{L_θ} decreased further. Because of the correlation between T_{L_θ} and pilot opinion between M_q of -1.0 and -3.0, one might expect then that pilot opinion would have continued to improve, but it actually deteriorated slightly at the high damping. Position loop lead T_{L_x} seems to explain this deterioration. Position loop lead remained constant at the lower values of M_q but increased notably at M_q of -5.0.

A decrease in information coupling between the pitch and position loops seems to account for the increase in T_{L_x} at M_q of -5.0. At the lower values of M_q , pitch attitude is responsive

TABLE 1.—COOPER PILOT RATING SYSTEM

| Operating conditions | Adjective rating | Numerical rating | Description | Primary mission accomplished | Can be landed |
|----------------------|------------------|------------------|---|------------------------------|---------------|
| Normal operation | Satisfactory | 1 | Excellent, includes optimum | Yes | Yes |
| | | 2 | Good, pleasant to fly | Yes | Yes |
| | | 3 | Satisfactory, but with some mildly unpleasant characteristics | Yes | Yes |
| Emergency operation | Unsatisfactory | 4 | Acceptable, but with unpleasant characteristics | Yes | Yes |
| | | 5 | Unacceptable for normal operation | Doubtful | Yes |
| | | 6 | Acceptable for emergency condition only ^a | Doubtful | Yes |
| No operation | Unacceptable | 7 | Unacceptable even for emergency condition ^a | No | Doubtful |
| | | 8 | Unacceptable—dangerous | No | No |
| | | 9 | Unacceptable—uncontrollable | No | No |
| | Catastrophic | 10 | Motions possibly violent enough to prevent pilot escape | No | No |

^a Failure of a stability augmenter.

TABLE 2.—VTOL AIRCRAFT OPEN-LOOP LONGITUDINAL TRANSFER FUNCTIONS

| General form of the transfer functions | Configurations | | | Factors of Δ_1 |
|---|----------------|-------|---------|-------------------------------|
| | M_q | X_u | $M_u g$ | |
| $\Delta_1 = s^3 - (M_q + X_u)s^2 + (M_q X_u)s + M_u g$ | -1.0 | -0.1 | 0.667 | $s+1.4, s-(0.14 \pm j0.68)$ |
| $\theta/\delta_P = \frac{M \delta_P (s - X_u)}{\Delta_1}$ | -3.0 | - .1 | .667 | $s+3.1, s+(0.013 \pm j0.47)$ |
| | -5.0 | - .1 | .667 | $s+5.0, s+(0.037 \pm j0.36)$ |
| $\theta/u_g = \frac{M_u s}{\Delta_1}$ | -3.0 | -0.1 | 0.0 | $s, s+3.0, s+0.10$ |
| | -3.0 | - .1 | .333 | $s+3.0, s+(0.031 \pm j0.33)$ |
| $x/\delta_P = \frac{-M_{\delta_P} g}{s \Delta_1}$ | -3.0 | - .1 | .667 | $s+3.1, s+(0.013 \pm j0.47)$ |
| | -3.0 | - .1 | 1.0 | $s+3.1, s-(0.0040 \pm j0.57)$ |
| $x/u_g = \frac{X_u (s^2 - M_q s) - M_u g}{s \Delta_1}$ | -3.0 | 0.0 | 0.667 | $s+3.0, s-(0.018 \pm j0.33)$ |
| | -3.0 | - .05 | .667 | $s+3.1, s-(0.011 \pm j0.47)$ |
| | -3.0 | - .1 | .667 | $s+3.1, s+(0.013 \pm j0.47)$ |
| | -3.0 | - .2 | .667 | $s+3.1, s+(0.062 \pm j0.46)$ |
| | -3.0 | - .3 | .667 | $s+3.1, s+(0.11 \pm j0.45)$ |

to turbulence and the pilot can anticipate being blown off in position by observing the pitch response. He can, therefore, develop control inputs to attenuate the induced position rates before they become large and thereby more easily maintain hovering accuracy. At M_q of -5.0, the pilot notices little pitch motion due to turbulence because the pitch response is heavily damped. Turbulence, acting through the drag parameter X_u , continues to introduce position rates but the pilot has no way to predict them. When the pilot cannot anticipate position disturbances through pitch response, he can maintain hovering accuracy only by responding more rapidly to the position rates he observes, that is, by supplying more T_{L_x} . Supplying more T_{L_x} demands increased concentration on position and faster response from the pilot and probably accounts for the deterioration in opinion.

Since the pilot could select optimum control sensitivity, his pitch-loop internal gain, K_{P_θ} , would also be expected to be optimum. Therefore, K_{P_θ} is not considered to be a factor in determining opinion. As shown in figure 5, K_{P_θ} increased almost linearly with increasing

M_q . For the unstable, highly oscillatory pitch dynamics at M_q of -1.0 the pilot used smaller stick inputs per unit error (smaller K_{P_θ}) in pitch than he did for the stable, well damped

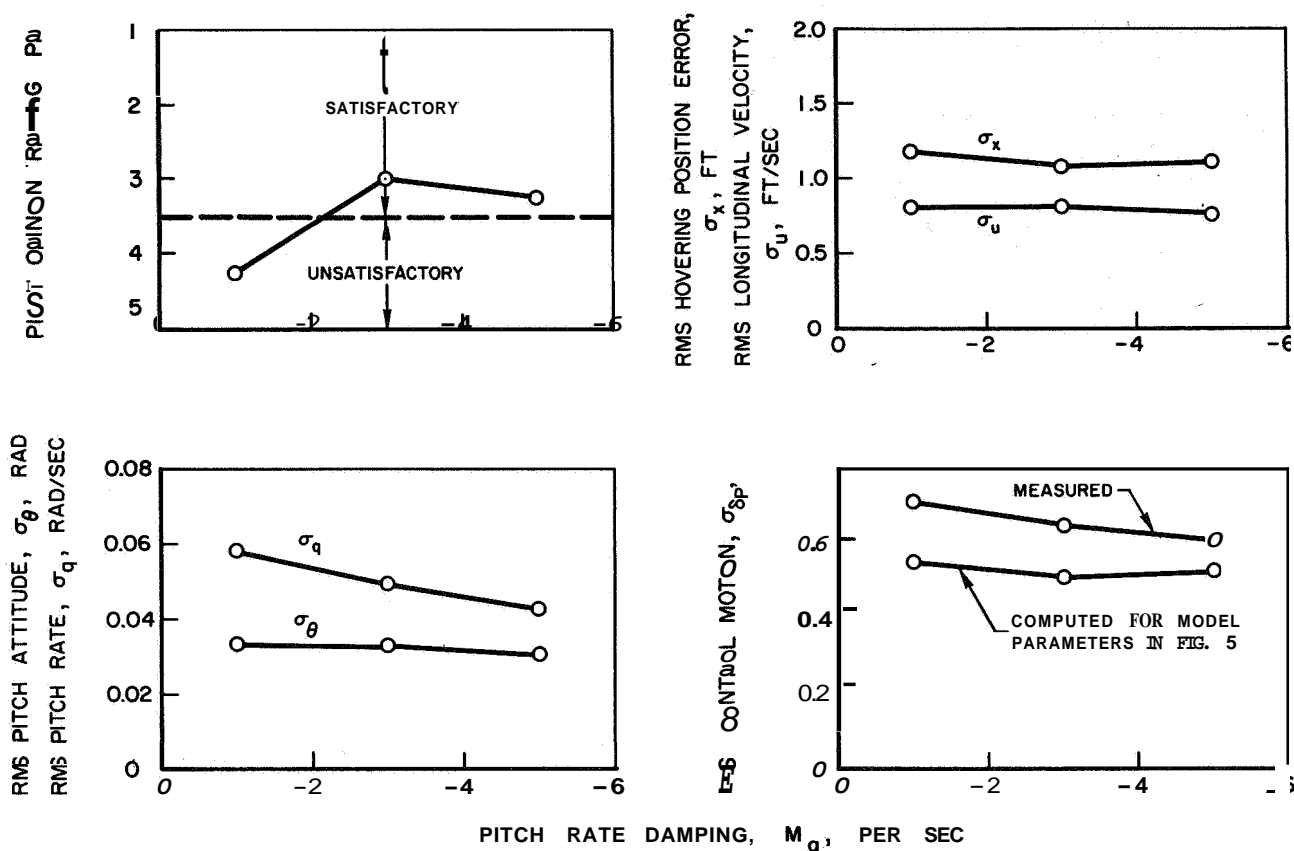


Figure 4.—Effect of pitch rate damping on pilot rms hovering performance. $M_{ug}=0.667$; $X_u=-0.1$; $\sigma_{ug}=5.14$; $\omega_B=0.314$; optimum $M_{\delta p}$.

dynamics at $M_q = -5.0$. Position-loop gain, K_{p_x} , remained nearly constant at M_q of -1.0 and -3.0 , but decreased at -5.0 . For these data, K_{p_x} does not appear to have any influence on pilot rating.

Although the fundamental characteristics of human-pilot control can be described by a linear model, the pilot also exhibits nonlinear behavior. He makes control motions when none are required, underestimates and overestimates the magnitude of control inputs necessary, etc. An indication of the pilot's nonlinearity is the difference between the measured rms control stick activity and the rms control activity computed using the linear pilot model and model parameters recovered for a given configuration. For the M_q variation, this difference is shown in figure 4.

Computed Closed-Loop Characteristics.—Bode plot loop-closure characteristics, crossover frequency ω_c , and phase margin (PM) for the M_q model results shown in figure 5 were as follows:

| M_q | ω_{c_θ} | PM $_\theta$ | ω_{c_x} | PM $_x$ |
|-------|---------------------|--------------|----------------|---------|
| -1.0 | 3.4 | 8 | 1.1 | 21 |
| -3.0 | 3.1 | 8 | 1.0 | 15 |
| -5.0 | 2.8 | 10 | 0.97 | 19 |

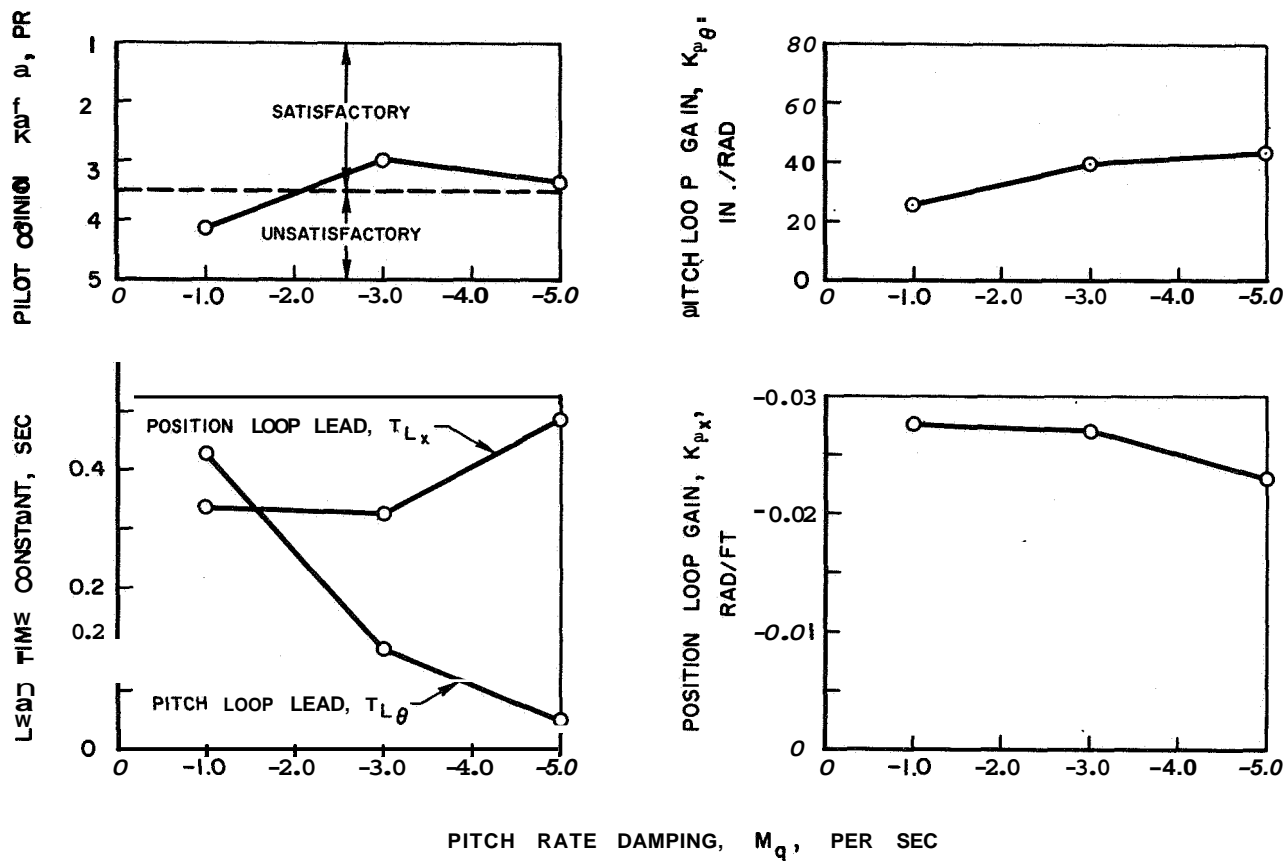


Figure 5.—Effect of pitch rate damping on pilot model parameters for hovering. $M_{ug}=0.667$; $X_u = -0.1$; $\sigma_{ug}=5.14$; $\omega_B=0.314$; optimum M_{δ_P} .

| M_q | x-loop slope, db/decade | θ -loop slope, db/decade |
|-------|-------------------------|---------------------------------|
| -1.0 | -32 | -34 |
| -3.0 | -32 | -38 |
| -5.0 | -31 | -35 |

The results of reference 7 indicated a consistent tendency of their operators to close the control loop with a -20 db/decade slope at crossover (single loop with K, K/s, and K/s² controlled elements). The higher slopes in this study probably result from the increased complexity of the controlled element (VTOL aircraft) dynamics. The pitch-loop slope is more negative than the position-loop slope because the pilots considered the 9-loop to be an intermediary in the primary task of controlling position. It appears that, in the process of controlling hovering position, the pilots closed the pitch loop more tightly than they would have if they were only controlling pitch. This tighter closure is reflected in the more negative slope.

RESULTS FOR VARIATIONS IN LONGITUDINAL SPEED-STABILITY PARAMETER, M_{ug}

Flight Simulator Data.—Optimum control sensitivities, pilot ratings, and rms hovering performance data were obtained for four values of speed-stability parameter, M_{ug} (0, 0.333, 0.667, and 1.0). The remaining aircraft stability derivatives and turbulence characteristics were held constant at their "common point" values when M_{ug} was varied. The open-loop aircraft transfer functions for these configurations are shown in table 2. The pilot selected the following values of optimum control sensitivity:

| M_{ug} | $M_{\delta P}$ |
|----------|----------------|
| 0 | 0.300 |
| .333 | .360 |
| .667 | .431 |
| 1.00 | .481 |

The pilot ratings and rms hovering performance data for these configurations are presented in figure 6.

The effect of M_{ug} on the measured rms values is shown in figure 6. Both σ_{θ} and σ_q increased with increasing M_{ug} . However, rms pitch rate increased more rapidly than rms pitch attitude. The measured values of σ_x and σ_u exhibited some random variations as M_{ug} increased, and there appeared to be no systematic trends with M_{ug} . Control-stick activity increased steadily with M_{ug} , reflecting the larger pitch disturbances which had to be controlled at higher M_{ug} . The variation in PR does not seem to correlate with any of the measured rms functions.

Computed Pilot Model Parameters.—The pilot model adaptive parameters computed from the rms results for variable M_{ug} are plotted in figure 7. For $M_{ug}=0$ the model results indicate that the pilot used a large T_{Lx} and almost no $T_{L\theta}$. There was no pitch-position

coupling, so he received no information from pitch disturbances to predict position rates, and therefore had to supply a large T_{Lx} to hover accurately. For $M_{ug}=0$, PR was on the unsatisfactory borderline.

As M_{ug} increased to 0.333, pitch response to turbulence became evident, $T_{L\theta}$ increased, and T_{Lx} decreased. Pilot rating improved for $M_{ug}=0.333$. The

pilot was aware of pitch disturbances due to turbulence although it was not annoying. This position acceleration information from the pitch disturbances seems to account for the drop in T_{Lx} and improvement in PR. As M_{ug} was increased further to 0.667, T_{Lx} became

smaller, $T_{L\theta}$ became larger, and PR further improved. Since the best PR was assigned at

M_{ug} of 0.667, the pilot apparently regarded this case as representing the best compromise between control activity required to control pitch disturbances at high M_{ug} and the amount of

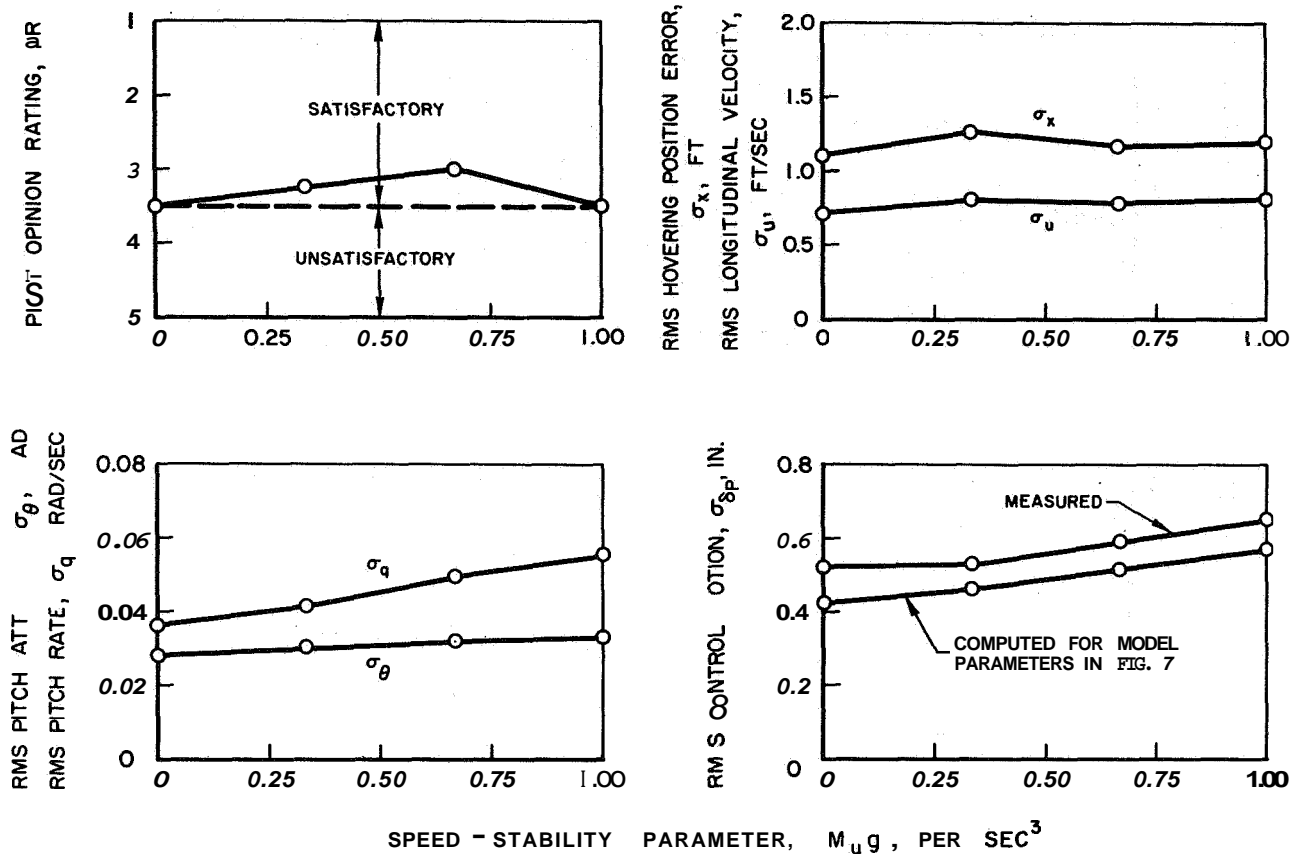


Figure 6.—Effect of longitudinal speed stability on pilot hovering performance. $X_u=0.1$; $M_q=-3.0$; $\sigma_{u_g}=5.14$; $\omega_B=0.314$; optimum M_{δ_P} .

T_{L_x} required at low M_{u_g} . At $M_{u_g}=1.0$, the activity and concentration required to control pitch disturbances became disagreeable to the pilot, and opinion deteriorated to the unsatisfactory borderline. This increased effort required to control pitch is reflected by the higher control sensitivity and control motion at $M_{u_g}=1.0$. There was no significant change in T_{L_x} between $M_{u_g}=0.667$ and 1.0 .

The variation in pitch-loop pilot gain K_{P_θ} with M_{u_g} is also shown in figure 7. It appears that K_{P_θ} was nearly invariant with M_{u_g} except that a slightly lower value is evident at $M_{u_g}=0$. It appears that the pilot's criterion for selecting M_{δ_P} as M_{u_g} varied was to keep his internal pitch-loop gain constant. The variation in the real pole location for the pitch dynamics was not as large for M_{u_g} changes as when M_q was varied (table II). This fixed position of the real pole may explain why K_{P_θ} was nearly constant for changes in M_{u_g} and not M_q .

Position-loop gain, K_{P_x} (fig. 7) increased as M_{u_g} became larger. It does not correlate with the trend in PR and, just as in the investigation of the effect of varying M_q , it appears that K_{P_x} did not significantly affect opinion.

Computed control stick activity for variable M_{u_g} is shown in figure 6 along with that measured from the human pilot. As discussed previously, the difference is an indication of pilot nonlinearity.

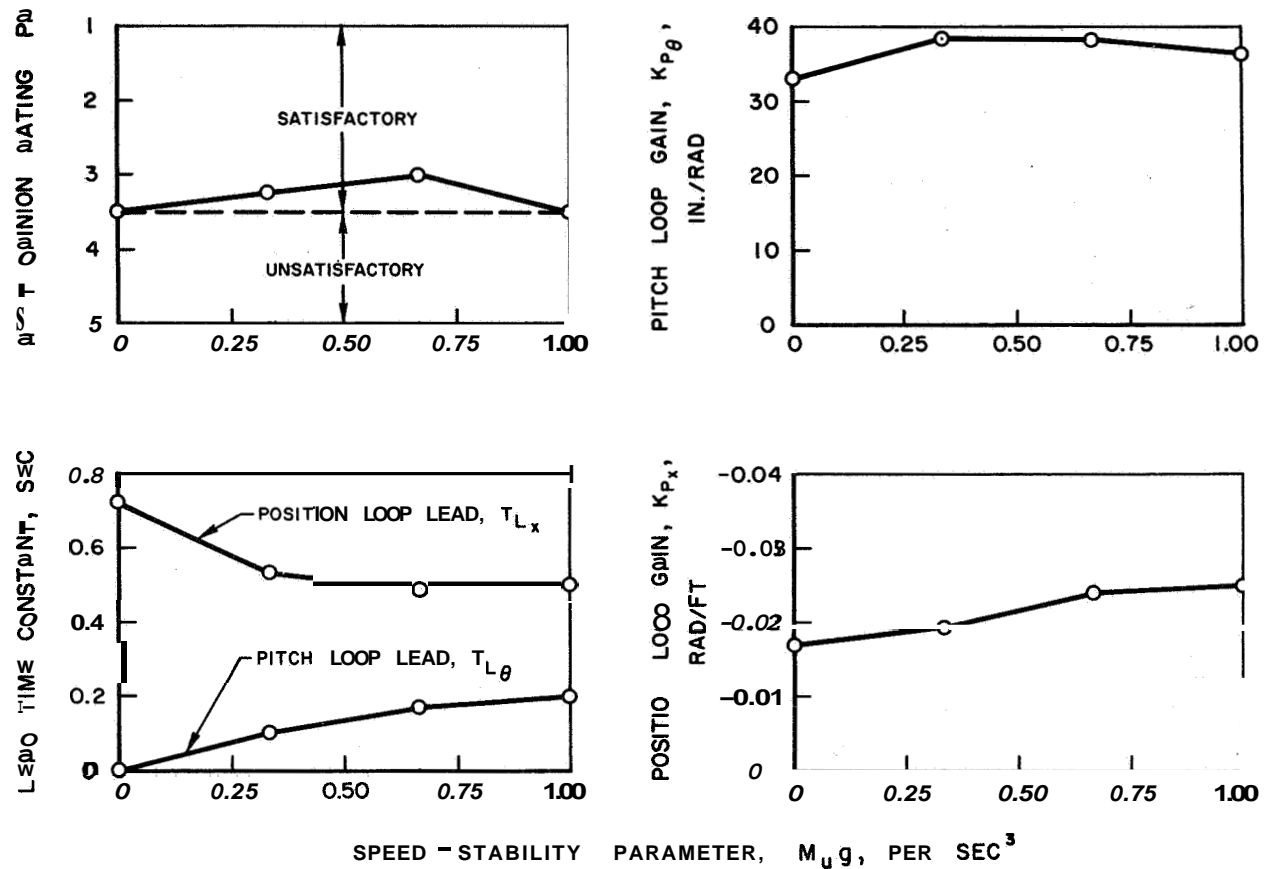


Figure 7.—Effect of longitudinal speed stability on pilot model parameters for hovering. $X_U = -0.1$; $M_q = -3.0$; $\sigma_u g = 5.14$; $\omega_B = 3.14$; optimum $M_\delta P'$

Computed Closed-Loop Characteristics.—Bode plot crossover frequencies ω_c and phase margins (PM) for the variable M_{ug} model results were:

| M_{ug} | $\omega_{c\theta}$ | PM_θ | ω_{cx} | PM_x |
|----------|--------------------|-------------|---------------|--------|
| 0 | 2.2 | 7 | 0.86 | 19 |
| .333 | 2.7 | 7 | .88 | 18 |
| .667 | 3.0 | 9 | .96 | 19 |
| 1.0 | 3.2 | 10 | .99 | 21 |

The variation in $\omega_{c\theta}$ was larger than that for ω_{cx} . It appears that the pilot was more concerned with maintaining consistent crossover frequency for the position loop than the pitch loop when M_{ug} was varied. This result is similar to that noted for variable M_q Bode plot data, and again it seems to reflect the pilot's desire to maintain constant hovering accuracy.

The slopes of the x-loop and θ -loop Bode magnitude plots at the crossover frequencies were calculated to be:

| M_{ug} | x-loop slope,
db/decade | 8-loop slope,
db/decade |
|----------|----------------------------|----------------------------|
| 0 | -28 | -33 |
| .333 | -32 | -37 |
| .667 | -32 | -37 |
| 1.0 | -32 | -37 |

Again, the slopes of both loops are considerably more negative than the -20 db/decade noted in reference 7, and the pitch-loop slope was more negative than the position-loop slope.

RESULTS FOR VARIATIONS IN LONGITUDINAL DRAG PARAMETER, X_u

Flight Simulator Data.—Optimum control sensitivities, pilot ratings, and rms hovering performance data were obtained for five values of drag parameter X_u (0, -0.05 , -0.1 , -0.2 , and -0.3). The remaining aircraft stability derivatives and turbulence characteristics were held constant at their "common point" values. Open-loop aircraft transfer functions for these configurations are shown in table 2. The pilot selected the following values of optimum control sensitivity:

| X_u | $M_{\delta P}$ |
|-------|----------------|
| 0 | 0.287 |
| - .05 | .420 |
| - .1 | .431 |
| - .2 | .469 |
| - .3 | .516 |

Pilot ratings and rms hovering performance data for these configurations are presented in figure 8.

The effect of X_u on measured rms results is shown in figure 8. These results indicate that hovering accuracy in turbulent air is much more dependent on X_u than M_{ug} or M_q (figs. 4 and 6). Rms hovering position error increased proportionately with increasing X_u , as did rms position rate. Rms pitch attitude and rms pitch rate also increased as X_u became larger. Rms control stick activity increased and pilot opinion deteriorated with increasing X_u . Measured rms results for a second pilot have been obtained at $X_u = -0.05$ and are also shown. The measured rms results for the two pilots are very similar as are the pilot model parameters computed from the rms results.

Computed Pilot Model Parameters.—The correlation between $T_{L\theta}$, T_{Lx} and pilot opinion for the variable M_q and M_{ug} results is not evident in the pilot model parameter results for variable X_u . As shown in figure 9, the steady deterioration in PR as X_u increased is accompanied by a marked decrease in T_{Lx} and a small reduction in $T_{L\theta}$. The reason for this reversal in the correlation between T_{Lx} and opinion is that the pilot was unwilling to make the large, rapid changes in pitch attitude which were required to supply more T_{Lx} as X_u became

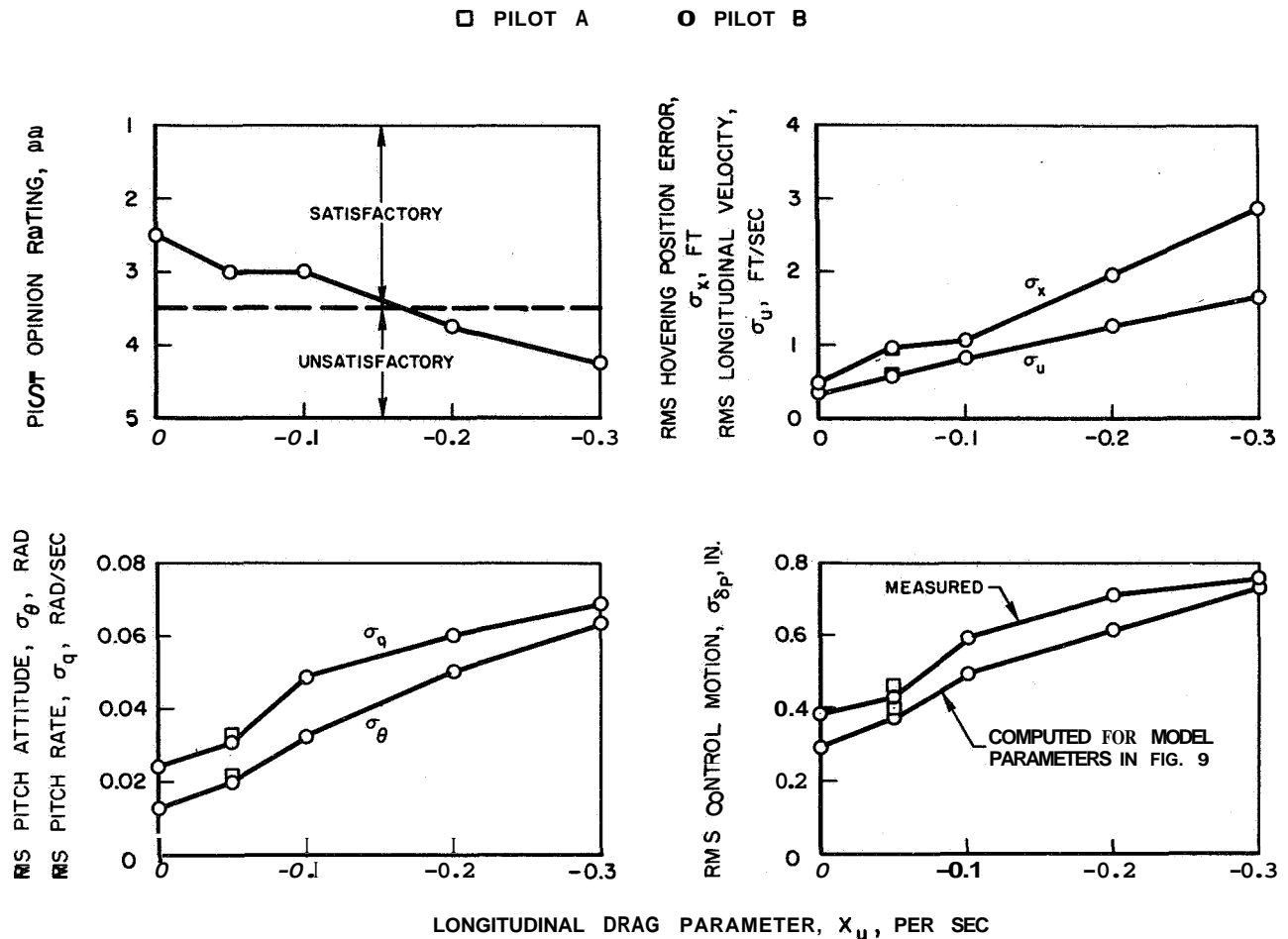


Figure 8.—Effect of longitudinal drag parameter on pilot rms hovering performance. $M_{Ug} = 0.667$; $M_{q} = -3.0$; $\sigma_{Ug} = 5.14$; $\omega_B = 0.314$; optimum $M_{\delta P}$.

larger. For the hovering task the pilot responds to position rates, that is, supplies T_{Lx} by changing pitch attitude (rotating the thrust vector). For small X_u , position disturbances caused by turbulence are correspondingly small and supplying T_{Lx} results in a low level of pitch attitude activity. If X_u is increased the position disturbances become larger and supplying the same T_{Lx} results in increased pitch activity (assuming the other pilot adaptive parameters remained fixed). An example of this is shown in the following table. Here the variation in analytical values of σ_θ , σ_q , σ_x , and σ_u for the task model are listed for four values of X_u . The pilot adaptive parameters and control sensitivity were held constant for this computation ($K_{P\theta} \cdot M_{\theta P} = 17.3$, $K_{P_x} = -0.0270$, $T_{L\theta} = 0.170$, and $T_{Lx} = 0.400$).

| X_u | σ_θ | σ_q | σ_x | σ_u |
|-------|-----------------|------------|------------|------------|
| 0 | 0.0145 | 0.0326 | 0.371 | 0.330 |
| - .1 | .0328 | .0492 | 1.07 | .803 |
| - .2 | .0499 | .0648 | 1.67 | 1.14 |
| - .3 | .0658 | .0794 | 2.22 | 1.41 |

As X_u was increased, σ_θ and σ_q increased proportionately reflecting the increased pitch rates and pitch attitude angles necessary to supply T_{Lx} . Note also that σ_x increased with increasing X_u and fixed pilot model parameters. The pilot must supply more T_{Lx} and K_{p_x} to maintain hovering accuracy as X_u is increased. Of course, this would result in an even faster increase in pitch activity with increased X_u . The computed pilot model results in figure 9 indicate that the pilot prefers to allow hovering accuracy to deteriorate, while keeping pitch activity within what he considers to be reasonable limits under these circumstances.

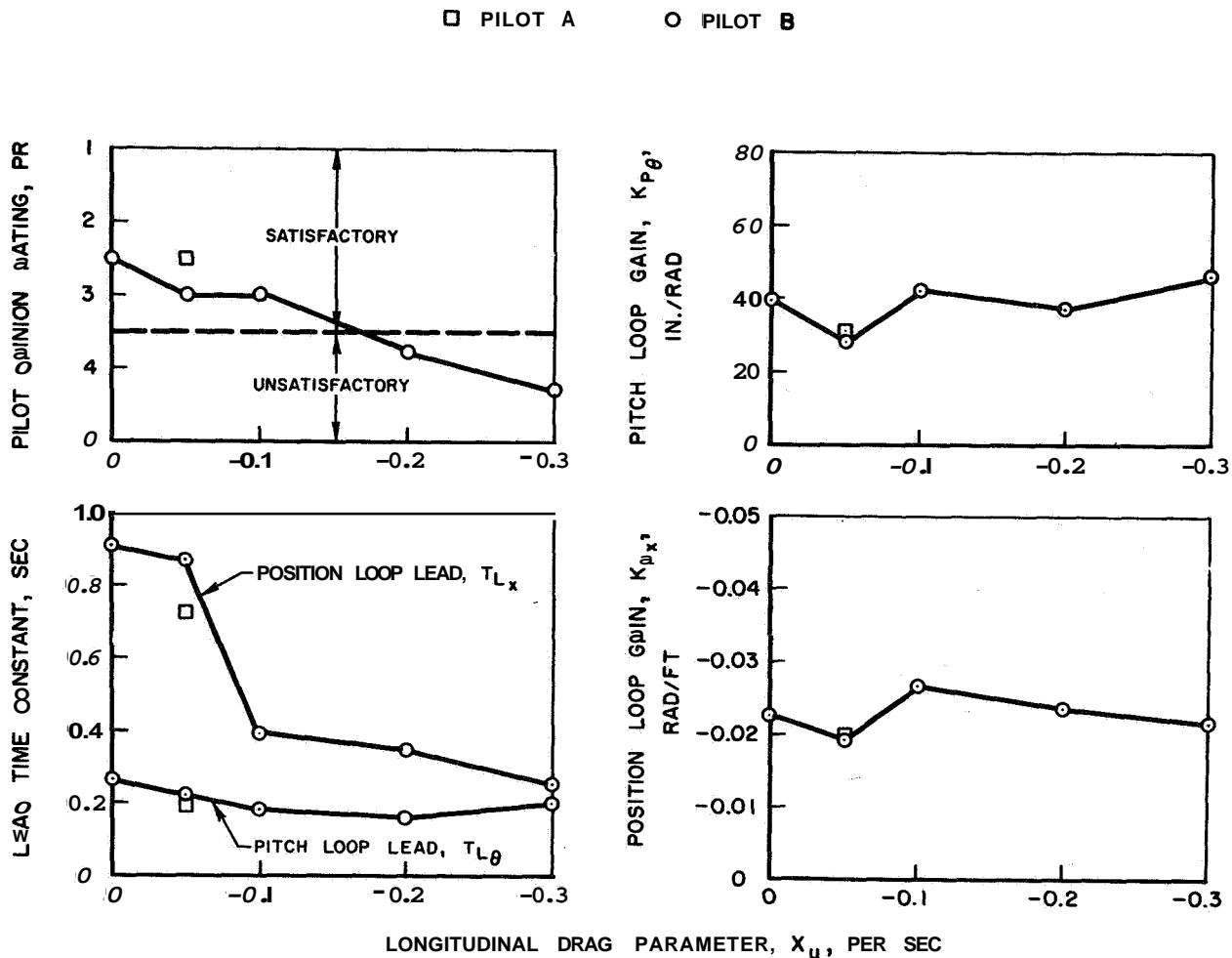


Figure 9.—Effect of longitudinal drag parameter on pilot model parameters for hovering.

$$M_{u_g}=0.667; M_q=-0.3.0; \sigma_{u_g}=5.14; \omega_B=0.314; \text{optimum } M_{\delta_P}.$$

Examining the values of T_{Lx} shown in figure 9 computed for each level of X_u shows that for $X_u=0$, T_{Lx} is largest; large position rates are not present since position rates only result from pitch attitude upsets. Large attitude changes are not required even to supply large T_{Lx} because the position rates are so small. For $X_u=-0.1$, T_{Lx} is less than its value for $X_u=0$. However, turbulence induced position rates are larger and of a higher frequency than for zero X_u . Supplying the same T_{Lx} which was used at $X_u=0$ would result in large pitch rates

and pitch angles if it were used for $X_u = -0.1$. This the pilot considers neither safe nor realistic and he therefore uses smaller T_{L_x} while allowing his hovering error to increase. The same trend in T_{L_x} continues as X_u increases to -0.2 and -0.3 . The rate of change of σ_x with X_u increased while the rate of change of σ_θ and σ_q with X_u decreased, indicating that the pilot prefers to limit the violence of pitch motion even though it means that he must also limit T_{L_x} and suffer a loss in hovering accuracy.

It was pointed out in the discussion of results for variable M_q that to control position accurately the pilot adapts a form of pitch-loop control characterized by low phase margin. This seems inconsistent with the preceding discussion, which indicates that the pilot sacrifices hovering accuracy in favor of pitch stability. The two results are not inconsistent. The pilot weighs performance against safety and realism. He is willing to destabilize pitch to some extent to achieve hovering accuracy, but not to the point where he begins to lose pitch control or develop pitch motions which could not be tolerated in an actual aircraft.

Pilot pitch loop gain, K_{p_θ} , shown in figure 9, appears to have been nearly independent of X_u . This was also the case with variations in M_{u_g} . The average K_{p_θ} for the M_{u_g} variation was 36 in./rad. The average K_{p_θ} for the X_u variation was 39 in./rad. Position-loop gain, K_{p_x} , shown in figure 9, again does not seem to correlate with opinion. No definite trend in K_{p_x} is apparent as opinion deteriorates with increasing X_u .

Computed control stick activity for variable X_u is shown in figure 8 along with that measured from the human pilot,

Computed Closed-Loop Characteristics.—Bode plot crossover frequencies ω_c and phase margins (PM) for the variable X_u model results were as follows:

| X_u | ω_{c_θ} | PM_θ | ω_{c_x} | PM_x |
|-------|---------------------|-------------|----------------|--------|
| 0 | 2.6 | 27.3 | 0.73 | 23.6 |
| - .05 | 2.6 | 22.2 | 1.05 | 27.9 |
| - .1 | 3.2 | 7.3 | 1.0 | 15.1 |
| - .2 | 3.1 | 6.2 | .92 | 19.8 |
| - .3 | 3.8 | 3.1 | .83 | 24.0 |

Pitch-loop crossover frequency increased with increasing X_u while PM_θ decreased. The decrease in PM_θ is due to the increase in position disturbances resulting from the larger X_u . As the difficulty in controlling hovering position increases, the pilot must drive the pitch loop towards instability in order to get the thrust vector motions necessary to control position rates. For $X_u=0$, PM_θ is nearly 30° , which is indicative of stable, reasonably well damped pitch-loop performance. For $X_u = -0.3$, PM_θ is only 3° , which is very near instability ($PM < 0$). Had the pilot not become concerned about pitch-loop stability at the higher X_u and limited his response to position rates, PM_θ might have been even smaller. Position-loop crossover frequency increases and PM_x decreases as X_u increases from 0 to -0.1 . As X_u increases from -0.1 to -0.3 , the trends in ω_{c_x} and PM_x reverse. It appears that the pilot was making some effort to control the increasingly larger position rates up to X_u of -0.1 by increasing the response (bandwidth) of the position loop. For X_u of -0.2 and -0.3 , the effects

of his decision to limit pitch motion became evident in the form of larger σ_x and σ_u , which implied more relaxed control of hovering position. This more relaxed control is evident, in terms of loop closure parameters, as decreasing ω_{c_x} and increasing PM_x .

The slopes of the x-loop and 8-loop Bode magnitude plots at the crossover frequencies for the variable X_u investigation are shown in the following table:

| X_u | x-loop slope,
db/decade | θ -loop slope,
db/decade |
|-------|----------------------------|------------------------------------|
| 0 | -31 | -32 |
| - .05 | -27 | -34 |
| - .1 | -33 | -38 |
| - .2 | -33 | -38 |
| - .3 | -35 | -38 |

These results are consistent with the slope characteristics for variable M_q and M_{u_g} . That is, the slopes are more negative than -20 db/decade and the 8-loop slope is more negative than the x-loop.

CONSISTENCY OF PILOT MODEL RESULTS FOR SIMILAR TEST CONFIGURATIONS.—

Of the 12 simulated VTOL configurations flown by the pilot, 3 were identical. That is, once during each part of the study when a different stability derivative was varied the "common point" configuration was evaluated. These common configurations were evaluated on dates separated by several weeks. Therefore, the consistency of the measured rms results and recovered parameters is an indication of the quality of the procedures used to measure data and recover pilot model parameters. The measured rms values and recovered model param-

| | σ_θ | σ_q | σ_x | σ_u | K_{P_θ} | K_{P_x} | T_{L_θ} | T_{L_x} |
|-------------------|-----------------|------------|------------|------------|----------------|-----------|----------------|-----------|
| M_q studies | 0.0330 | 0.0495 | 1.08 | 0.807 | 39.7 | -0.0270 | 0.168 | 0.405 |
| M_{u_g} studies | .0320 | .0498 | 1.17 | .794 | 37.5 | - .0240 | .162 | .491 |
| X_u studies | .0323 | .0484 | 1.08 | .803 | 42.0 | - .0268 | .175 | .389 |

With the possible exception of an increase in T_{L_x} for the "common point," from the M_{u_g} studies the results are quite consistent.

CONCLUDING REMARKS

The investigation reported in this paper has provided a considerable amount of information regarding a pilot's performance of a coupled, two-degree-of-freedom hovering task and the correlation of pilot model parameters with pilot opinion ratings and optimum control sensitivities. It can be concluded that quasi-linear pilot models and servoanalysis techniques are valuable tools for determining pilot control characteristics and interpreting pilot behavior.

However, additional exploratory investigations are required before these tools are refined sufficiently to be employed in the prediction of PR and optimum $M_{\delta P}$. These studies are being continued at the Research Laboratories. Pilot model parameters are being computed from the rms results measured for a second pilot who took part in the study. Continuing efforts are also being made to refine the model and explore further the implications of the model results.

REFERENCES

1. Salmirs, S.; and Tapscott, R. J.: The Effects of Various Combinations of Damping and Control Power on Helicopter Handling Qualities During Both Instrument and Visual Flight. NASA TN D-58, 1959.
2. Faye, A. E., Jr.: Attitude Control Requirements for Hovering Determined Through the Use of a Piloted Flight Simulator. NASA TN D-792, 1961.
3. Madden, J.; Kroll, J.; and Neil, D.: A Study of VTOL/STOL Flying Qualities Requirements. Rep. 203-917001, Bell Aerospace Corp., Aug. 1960.
4. Isca, J. A.; and Patierno, J.: Instrument Flight Simulator Study of the VTOL Controllability-Control Power Relationship. IAS paper 61-118-1812, June 1961.
5. Seckel, E.; Traybar, J. J.; and Miller, G. E.: Longitudinal Handling Qualities for Hovering. Proc. Eighteenth Annual Forum of American Helicopter Society, May 1962.
6. Miller, D. P.; and Clark, J. W.: Research on VTOL Aircraft Handling Qualities Criteria. J. Aircraft, vol. 2, no. 3, May 1965.
7. McRuer, D. T.; Graham, D.; Krendell, E. S.; and Reisener, W.: Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15, July 1965.
8. McRuer, D. T.; and Krendell, E. S.: Dynamic Response of Human Operators. WADC-TR-56-524, Oct. 1957.
9. Stapleford, R. L.; McRuer, D. T.; and Magdeleno, R.: Pilot Describing Function Measurements in a Multiloop Task. Second Annual NASA-University Conference on Manual Control, NASA SP-128, 1967, pp. 181-204.
10. Stapleford, R. L., et al: An Analytical Study of V/STOL Handling Qualities in Hover and Transition. AFFDL-TR-65-73, Oct. 1965.
11. Newton, G. C., Jr.; Gould, L. G.; and Kaiser, J. F.: Analytical Design of Linear Feedback Controls. John Wiley and Sons, Inc., 1957.

33. Man-Machine System Equalization Determination and its Implementation

C. L. Tipton
U.S. Naval Research Laboratory

This paper presents a simplified and direct method of system synthesis of higher order, man-machine control loops. A profile of system feedback is derived which produces a previously selected standard form of closed-loop step response. Details of this design procedure are outlined and, by way of example, the equalization for an eighth-order aircraft vertical-rate control loop is determined. Further, through an implementation procedure, the synthesized equalization formulation is executed in terms of the given system structure. Finally, a man-machine experiment compares the conventional and equalized control-loop configurations. The results show that the derived equalization effects a 90-percent reduction in both system error and the amount of control manipulation.

Higher order control loops with lag characteristics are a common feature of current man-machine-system configurations. Accurate and stable system performance is dependent in large measure on the appropriate equalization of the complex dynamics incurred in such loops. And in a manual mode of system operation it is necessary that this equalization be provided by the human controller. The human pilot, helmsman, or operator, however, does not optimally equalize higher order system dynamics. He can, in fact, become oscillatory and unstable.

The technique of including system feedback terms in an operator's error display is one design approach which alleviates this control task difficulty and significantly improves system performance (ref. 1). In this method of system compensation, selected derivative functions are summed with the output of the system parameter to be controlled. This composite feedback quantity is then compared with a command input, and the difference is presented to the operator as a display of error which must be minimized by compensating motions of his control (fig. 1). Although the system feedback loops to an operator's display are external to his machine proper, they are nonetheless effective in the modification of the man-machine-system transfer function. These loops constitute a form of "parallel feedback compensation" and permit the designer to manipulate the response characteristics of an existing system.

This paper presents a method for the direct determination and implementation of that profile of feedback terms specific to the optimal damping or equalization of any of a very large population of follower-type man-machine control systems. In general, the steps included are (1) the formulation of a linear mathematical (Laplacian) model of the system and the control loops of interest, (2) the selection of an "optimal" form of control loop (transient) response, (3) the mathematical determination of that profile of feedback compensation which produces the selected transient response, (4) the diagrammatic implementation in the given system's structure of the mathematically synthesized equalization formulation, and (5) the laboratory simulation and exercise of the given man-machine-system configuration to substantiate the relative efficacy of the prescribed treatment. To present this methodology in a practical

context a sample higher-order man-machine control loop is evaluated and experimentally validated in the discussion.

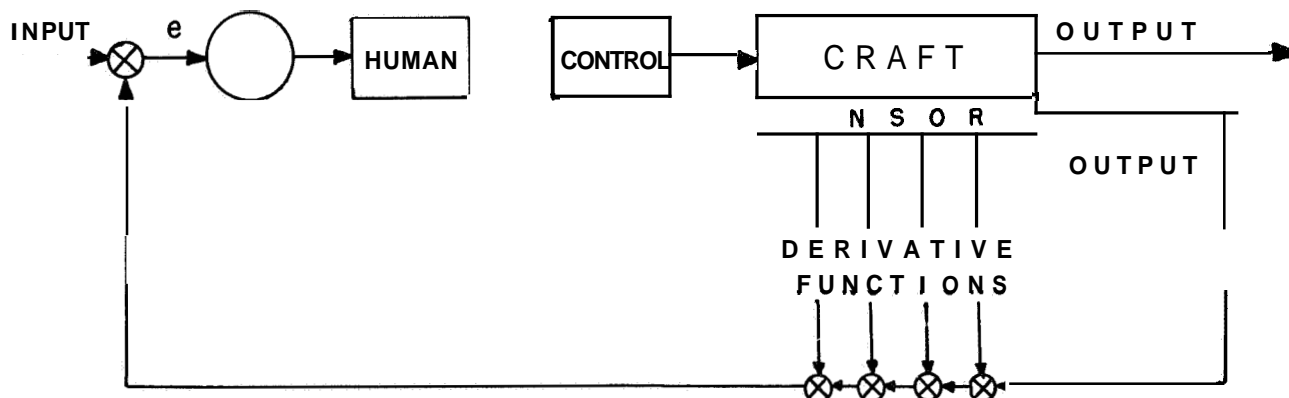


Figure 1.—General man-machine system control loop with compensating feedback.

SYSTEM DESCRIPTION

The study and treatment of a man-machine control loop is predicated on the derivation of a manipulatable mathematical model of the system dynamics and influences within which the given control loop will function. For illustrative purposes, consider it desirable to optimize the manual control of an aircraft's rate-of-descent in the landing mode of flight using throttle only. A block diagram of the man-machine-system configuration of a high-performance fixed-wing aircraft at a landing speed of 150 knots is presented in figure 2. The rate-of-descent or altitude-rate control loop as extracted from the general system description is presented as a block diagram in figure 3. As shown, the dynamic-system relevant description of the designated man-machine control loop is given by an eighth-order closed-loop transfer function.

These descriptive techniques, in the main, employ Laplace conventions (refs. 2 to 5). Such transforms are available from equipment manufacturers and/or linear differential equations approximating specific system elements and operations. One noteworthy exception occurs, however, with the inclusion of a human element. In the present example, the open-loop human/display/control assembly has been characterized as a low-gain integrator having the transform of $2.5/s$. This approximation was obtained by the simple expedient of observing the operator's step response in a closed-loop system containing a unity-gain display and a light-control level, that is, no dynamics. The rationalization for characterizing this controller assembly from its performance in a position-type system stems from the fact that such a condition is obtained in a fully equalized system; that is, the operator does not need to compensate any system dynamics and can operate as a simple servo element. The use of a simple control lever was an artifice employed for convenience in the later simulation of this example system. In practice, the Laplace transform describing the controller assembly would be obtained by a laboratory experiment using the gain and loads imposed by the display and control specific to the given vehicle.

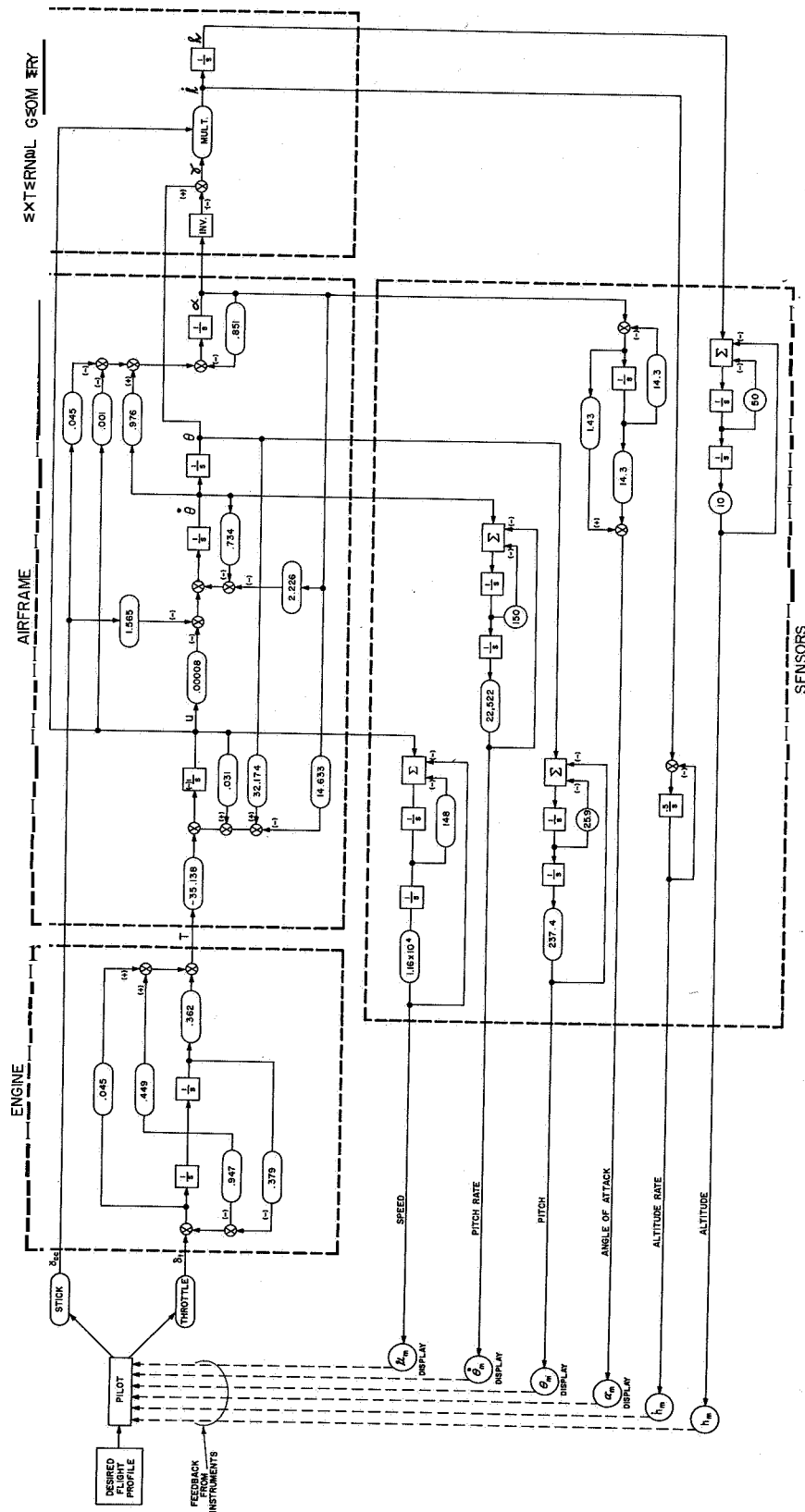


Figure 2.—A linearized block diagram of the man-aircraft system structure (longitudinal plane).

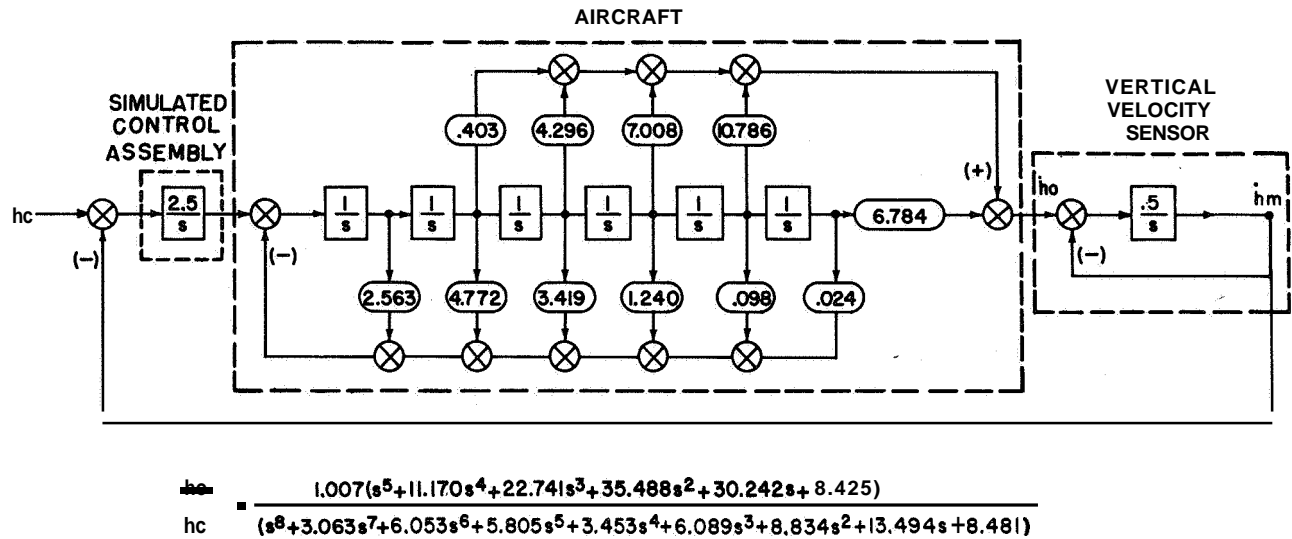


Figure 3.—Block diagram of the vertical rate man-machine control loop.

SYSTEM EQUALIZATION

A common index of the quality of system performance—and the quality of system compensation—is the time-amplitude character of a system's response to a step input. Several categorizations of system design exist which specify those standard form, unit-numerator transfer functions that produce a select form of transient response (refs. 6 and 7). Examples of such standard form transfer functions for systems of given order are provided by the ITAE, Butterworth, and binomial formulations. The ITAE (Integral of Time-multiplied Absolute-value of Error) formulation shown in figure 4 produces step responses characterized by a rapid-rise time and a quickly damped overshoot. The Butterworth formulations shown in figure 5 provide similar but more regular responses as a function of system order.

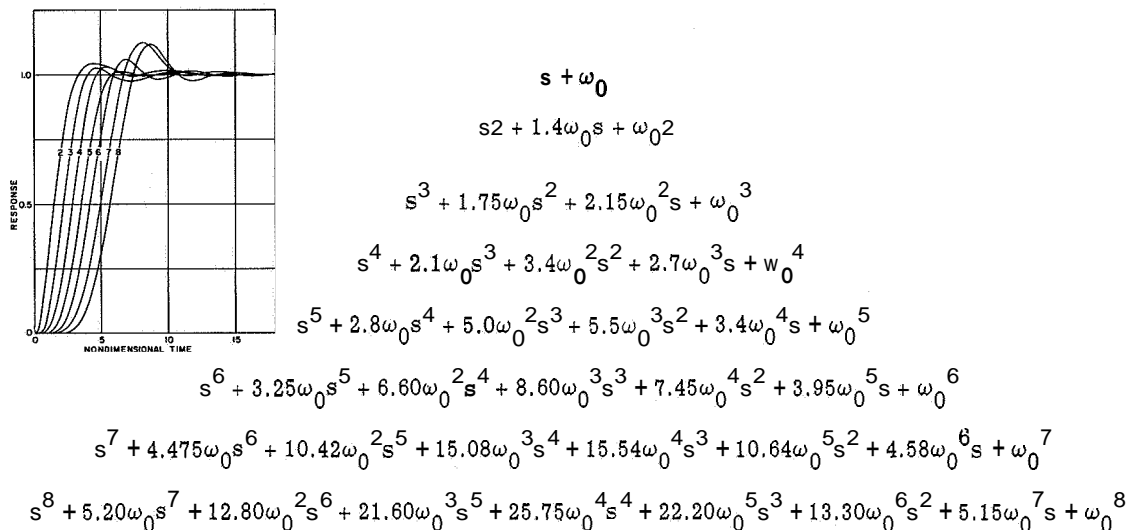


Figure 4.—The ITAE standard denominator forms.

The binomial category of standard transforms shown in figure 6 produces responses of a singularly stable form in a slower but asymptotic approach to maximum amplitude.

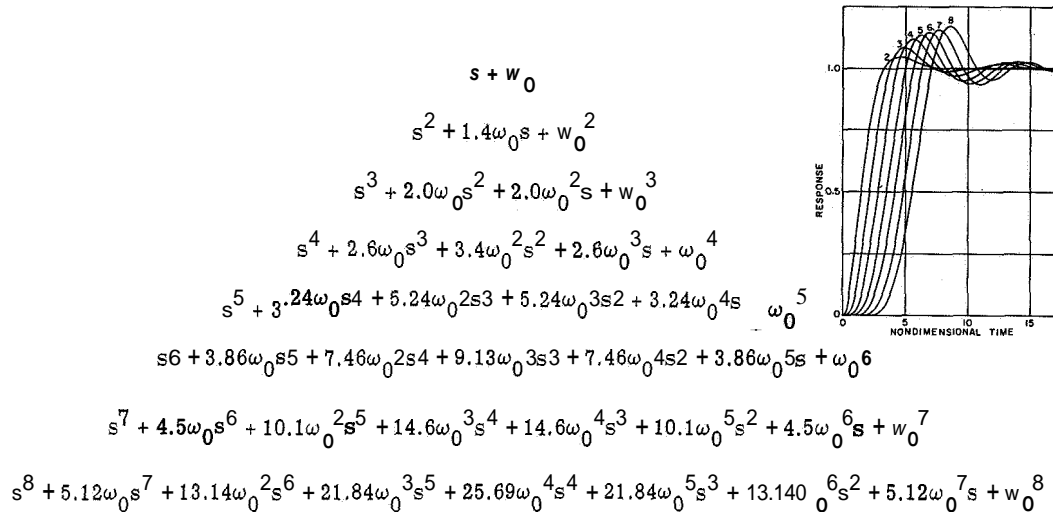


Figure 5.—The Butterworth standard denominator forms.

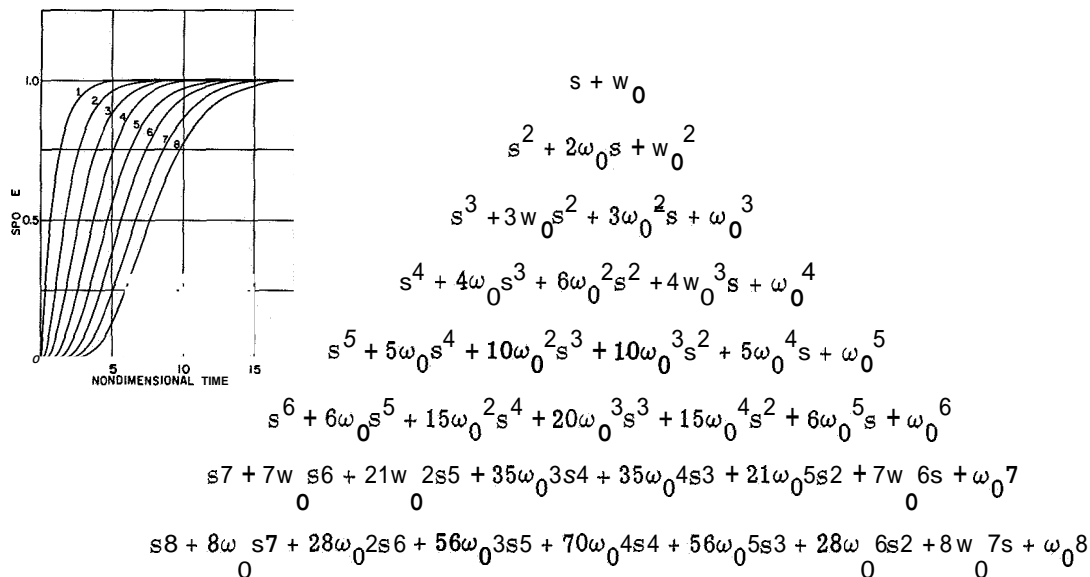


Figure 6.—The binomial standard denominator forms.

The ITAE, Butterworth, and binomial standard forms are employed here in determining that closed-loop denominator expression (feedback) which optimizes the step-response of a given control loop. In brief, an optimal profile of feedback terms is obtained from a simple multiplicative combination of a given loop's feedforward dynamics (numerator polynomial) and a selected standard form of an appropriate order. The resulting expression provides a

complete form of "cancellation compensation," that is, in a pole-zero configuration, closed-loop poles are created which compensate for all closed-loop zeros. The remainder of system poles are located in compliance with the criterion implicit in the selected standard form. The consequence of this operation is a system synthesis wherein a non-unit numerator, higher order control loop is made equivalent to a unit numerator, lower order, optimally responding standard form.

This process of optimizing or "standardizing" a given closed-loop transfer function is accomplished by a few simple and direct mathematical translations. The ITAE, Butterworth, and binomial standard form have the general expression

$$\frac{\theta_o}{\theta_i} = \frac{K}{R^p} = \frac{K}{s^p + (a_1 \omega_o) s^{p-1} + (a_2 \omega_o^2) s^{p-2} + \dots + \omega_o^p} \quad (1)$$

where :

θ_i input

θ_o output

s the Laplace operator

p the system order

K the system's feedforward gain

ω_o the system's natural angular frequency

a_1, a_2, \dots coefficients specified by the given standard (ITAE, Butterworth, binomial)

Man-machine control systems occurring in practice, however, contain feedforward dynamics giving rise to transfer functions of the general form

$$\frac{\delta_o}{\delta_i} = \frac{N^n}{D^m} = \frac{KN^n}{D^m} = \frac{K(s^n + k_1 s^{n-1} + k_2 s^{n-2} + \dots + k_n)}{s^m + c_1 s^{m-1} + c_2 s^{m-2} + \dots + c_m} \quad (2)$$

where :

δ_i input

δ_o output

s the Laplace operator

N^n the system's feedforward dynamics

K the system's given forward gain as specified by that factored constant which produces a unity coefficient in the s^n term of N^n

- D^m the system's given feedback dynamics
- k the system's given feedforward coefficients
- c the system's given feedback coefficients

A new, optimizing denominator D^m which will standardize the step response of the above generalized control loop is formulated from (1) the given system's feedforward dynamics N^n and (2) that standard form R^p which, in multiplicative combination with N^n , will produce a polynomial of the given system order m or

$$D^m = (N^n (R^p)) \quad (3)$$

The order of the standard form to be employed is simply

$$R^p = R^{m-n} \quad (4)$$

Thus, the given system's reconstructed closed-loop transfer function can be expressed as

$$\frac{\delta_o}{\delta_i} = \frac{K(N^n)}{D^m} = \frac{K(N^n)}{(R^{m-n})(N^n)} \quad (5)$$

from which there derives the treated system's equivalent lower order, unit numerator standard form

$$\frac{\delta_o}{\delta_i} = \frac{K}{R^{m-n}} \quad (6)$$

It remains only to determine the coefficients of the standard form expression R^{m-n} to specify fully the system optimizing expression D^m . These coefficients are implicit in the ITAE, Butterworth, and binomial formulations, respectively, and are a function of the system's natural angular frequency ω_o (eq. (1)). This frequency factor is obtained directly when ω_o is equated to the given system's forward gain K such that the system manifests a steady-state gain of unity. Such a condition is obtained when

$$\omega_o^{m-n} = K \quad (7)$$

With this derivation, the coefficients of R^{m-n} are determined and a given system's optimizing feedback is fully specified by

$$\begin{aligned} \frac{\delta_o}{\delta_i} &= \frac{K(N^n)}{D^m} = \frac{K(N^n)}{(R^{m-n})(N^n)} \\ &= \frac{K(s^n + k_1 s^{n-1} + k_2 s^{n-2} + \dots + k)}{(s^p + a_1 \omega_o s^{p-1} + a_2 \omega_o^2 s^{p-2} + \dots + \omega_o^p)(s^n + k_1 s^{n-1} + k_2 s^{n-2} + \dots + k)} \end{aligned} \quad (8)$$

Finally, the necessary feedback treatment or gain adjustments is given by the differences between the original system denominator D^m and the optimal form D^{im} or

$$\text{Gain adjustments} = GA = D^m - D^{im} \quad (9)$$

The results of this subtraction specify the negative and/or positive amount of any and each feedback term which must be inserted to obtain the selected form of step response. And, as stated earlier, this response optimization can be effected through the inclusion of the G_Δ terms in the composite feedback to an operator's display of system error.

A more specific example of these procedures and their influence is shown by the optimization of the vertical rate control loop using the Binomial standard. This example man-machine loop (fig. 3) has the closed-loop transform

$$\frac{\dot{h}_o}{\dot{h}_c} = K \frac{(N^n)}{D^m} = \frac{1.007(s^5 + 11.170s^4 + 22.741s^3 + 35.488s^2 + 30.242s + 8.425)}{s^8 + 3.063s^7 + 6.054s^6 + 5.805s^5 + 3.453s^4 + 6.089s^3 + 8.834s^2 + 13.494s + 8.481} \quad (10)$$

where

\dot{h}_c command input

\dot{h}_o system output

K that constant which factored from the feedforward (numerator) expression produces a unity coefficient in the highest order term s^n of N^n

The optimal closed-loop denominator (D^{im}) for this system is created from (eq. 8) the feed-forward dynamics (N^n) and the Binomial standard form R^{m-n} or:

$$D^{im} = (R^{8-5}) (N^5) = (R^3) (s^5 + 11.170s^4 + 22.741s^3 + 35.488s^2 + 30.242s + 8.425) \quad (11)$$

Referring to the binomial table of transforms, the third-order form requires that

$$R^3 = s^3 + 3\omega_o s^2 + 3\omega_o^2 s + \omega_o^3 \quad (12)$$

and from equation (7), the frequency factor ω_o is a function of K such that

$$\omega_o^{m-n} = K$$

whereby ,

$$\omega_o^3 = 1.007 \quad (13)$$

Thus , the eighth-order profile of feedback terms which will equalize to the binomial standard is given by

$$\begin{aligned} D^{1m} &= (s^3 + 3.007s^2 + 3.013s + 1.007)(s^5 + 11.170s^4 + 22.741s^3 + 35.488s^2 + 30.242s + 8.425) \\ &= s^8 + 14.177s^7 + 59.340s^6 + 138.528s^5 + 216.711s^4 + 229.178s^3 + 152.180s^2 + 55.828s + 8.481 \end{aligned} \quad (14)$$

Finally, the required modification of feedback terms to effect this form of system compensation is specified by

$$G_{\Delta} = D^m - D'^m$$

$$= -11.114063s^7 - 53.28606s^6 - 132.723088s^5 - 213.258401s^4 - 223.089417s^3 - 143.346225s^2 - 42.334158s \quad (15)$$

The analog diagram of this eighth order man-machine control loop with the derived equalization inserted as parallel feedback loops to the controller's display is presented in figure 7. The system is operable through either the human-control assembly or the analog (2.5/s) representation. Figure 8 shows the transient response of the analog mode and five consecutive responses using a human operator. The binomial form of these system responses attests to the efficacy of the equalization. Further, the similarity of all responses attest to the appropriateness of the empirical approximation of the man/display/control assembly for performing a system equalization analysis.

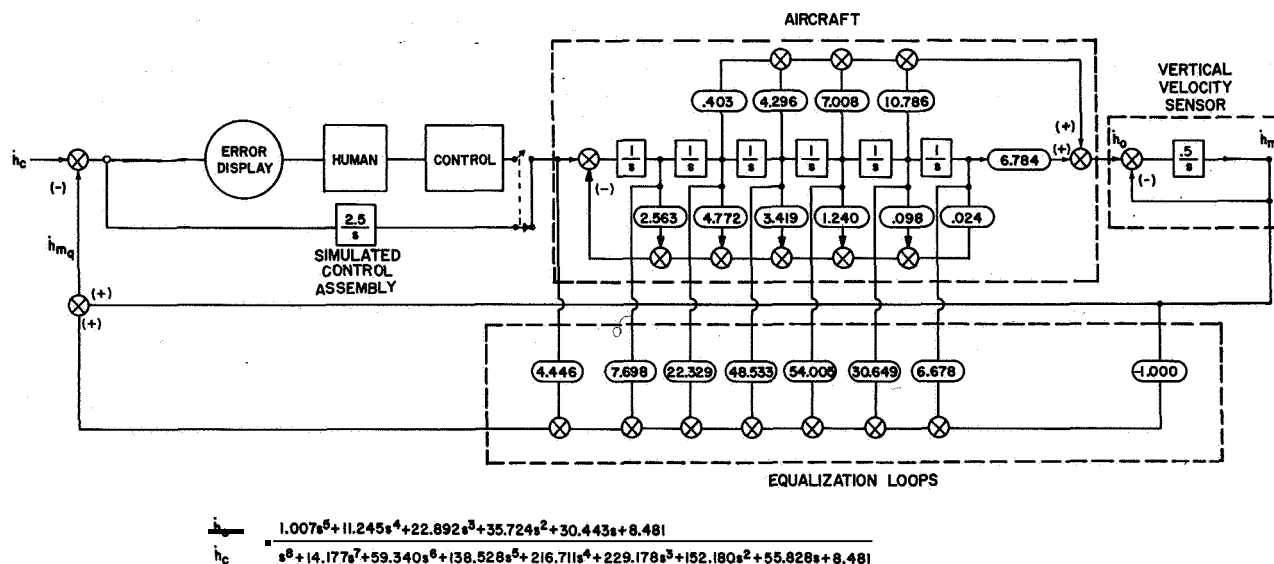


Figure 7.—Block diagram of the equalized vertical rate man-machine control loop.

That the derived equalization is a form of "cancellation compensation" is evidenced by the similarity of the loop's transient response to that of a third-order, unit-numerator analog. This fact is further confirmed by the s-plane plot of the equalized closed-loop pole-zero configuration presented in figure 9. The plot shows the cancellation of all system zeros by the superimposing of poles and the resulting system behavior described by the remaining third-order configuration of poles. The remaining poles are located as a function of the binomial standard form incorporating the system's forward gain. This example illustrates the foregoing method to be a simplified pole zero synthesis, that is, the direct derivation of a pole-zero configuration which produces a man-machine-system transform having a previously selected (standard) form of transient response.

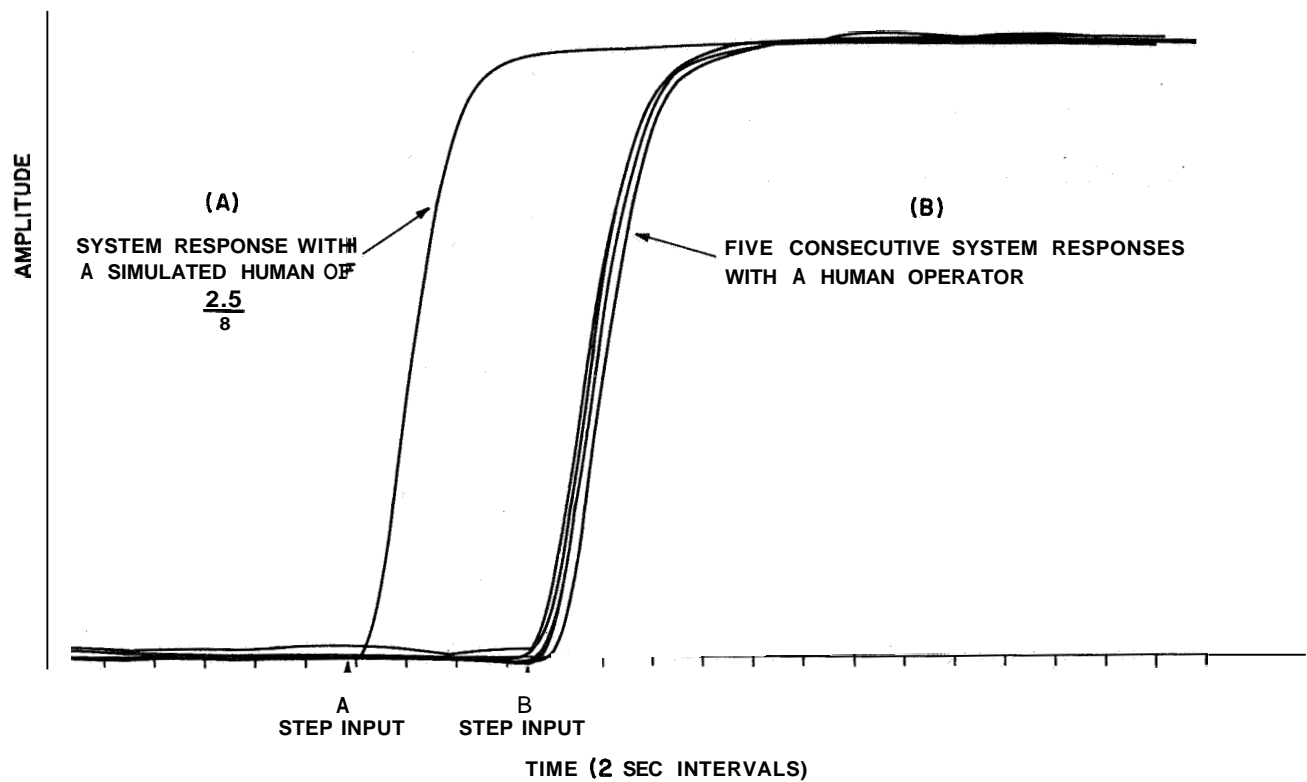
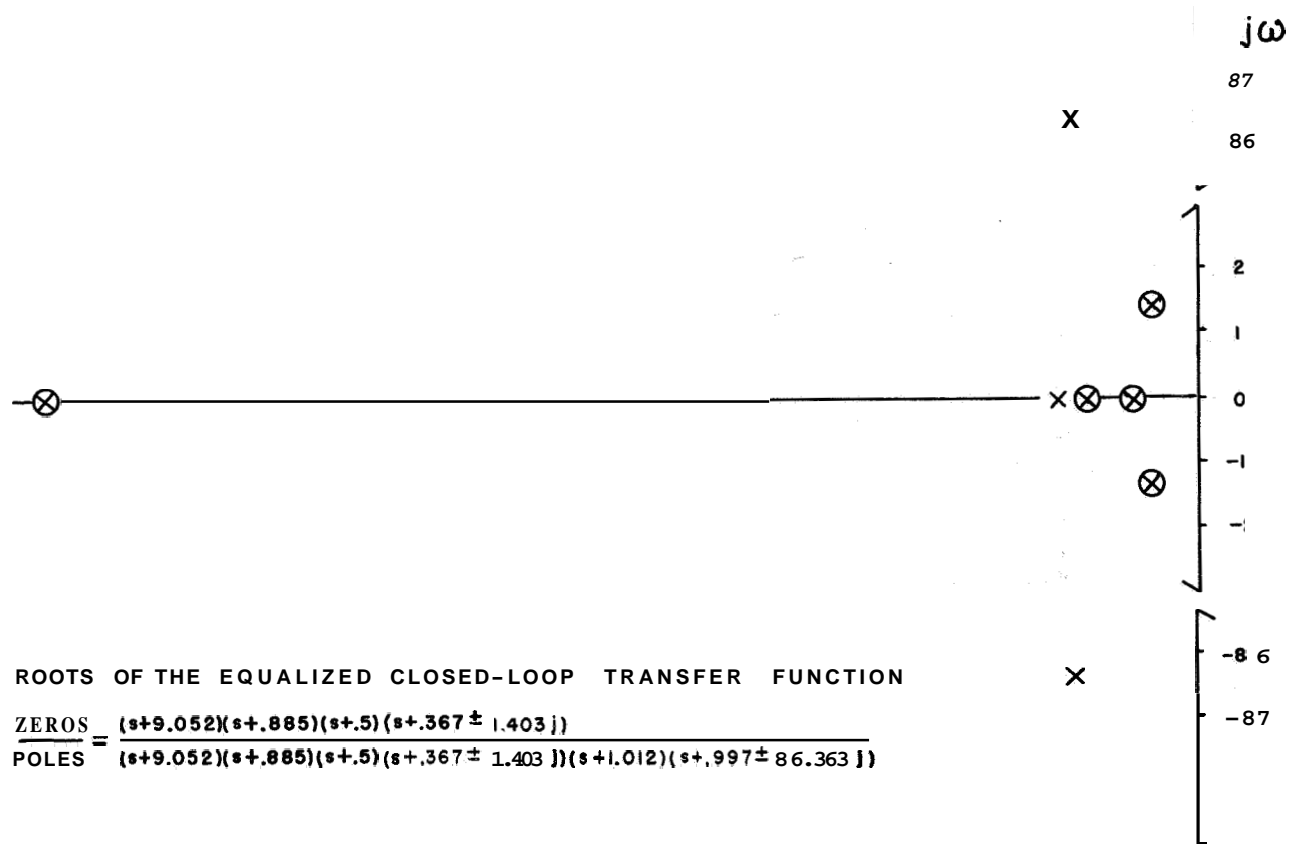


Figure 8.—The step response of the vertical rate control loop using an analog controller assembly of $2.5/s$ as compared with five consecutive step responses using a human operator.



EQUALIZATION IMPLEMENTATION

The foregoing method of system synthesis derives an optimized closed-loop system denominator. This polynomial specifies a profile of feedback terms which are direct derivative functions of the system output. In addition, the analog set-up of this mathematical system synthesis makes these derivative functions directly available for simulation and test purposes. This is not the case in an actual system. Simple derivative functions are not directly available from real system sensors. Further, an actual system seldom includes that number and location of sensors commensurate with the order and configuration of the system. Put succinctly, the real system usually contains an insufficient number of sensors and their outputs are complex functions of the system's dynamics. Thus, the very practical problem arises of implementing the optimal, but synthesized, closed-loop denominator in a realistic system.

This problem is readily handled by reformulating the system description in terms of the actual system structure containing parallel feedback loops of unknown gains. The selection of these unknown gains in terms of the optimal denominator form D' directly prescribes the real system treatment. The vertical-velocity control-loop structure is shown in figure 10. A parallel feedback loop has been added from the output of each integrating element in the system. All of these loops except one, the g_2 loop, are from existing or immediately available outputs. The g_2 loop refers to a figurative source sensing a function of engine thrust f_T . The other designations $g_{1,3,4,5,6,7,8}$ identify the outputs throttle position, thrust, speed, pitch rate, pitch, angle-of-attack, and the prime output, the measured vertical rate \dot{h}_m . As stated earlier, the throttle control function of unity was established for convenience in simulation. Transfer functions of unity have been assumed for all sensors except the barometric sensor which has a time constant of approximately 2 seconds. Although these assumptions of unity transforms are an analytical convenience, the included real sensors possess a rapidity and stability of response which warrants the disregard of their actual transfer functions.

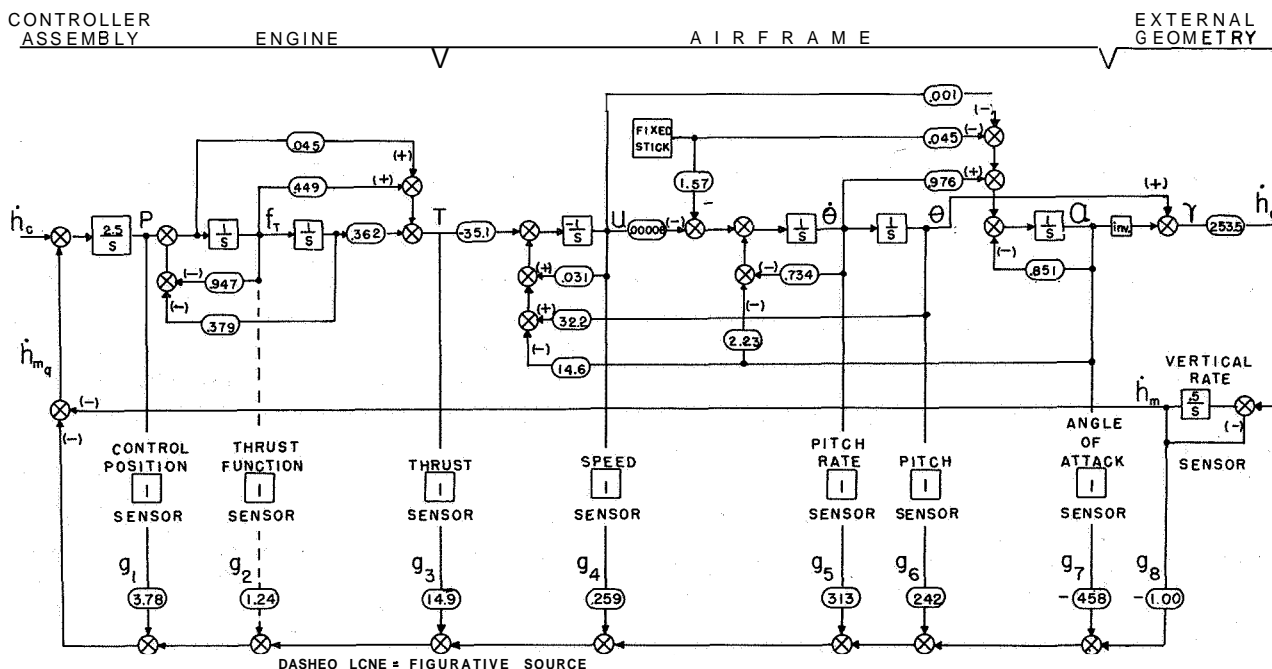


Figure 10.—The vertical-velocity control loop configured for the implementation of a derived equalization in terms of the man-machine system structure.

The closed-loop transfer function of this configuration is

$$\frac{h_o}{h_c} = \frac{N(s + .5)}{D(1 + g_1 P + g_2 f_T + g_3 T + g_4 U + g_5 \dot{\theta} + g_6 \ddot{\theta} + g_7 \alpha)(s + 0.5) + g_8 0.5N + 0.5N} \quad (16)$$

where

N open loop man-machine numerator

D open loop man-machine denominator

The closed-loop expression of this untreated system produces an eighth-order feedback polynomial whose coefficients are expressed in the eight variables g_1 to g_8 or

$$\begin{aligned} h_c = & s^8 \\ & + s^7(3.063 + 2.5g_1 + .1125g_3) \\ & + s^6(6.0535 + 7.6575g_1 + 2.5g_2 + 1.3606g_3 + 3.971g_4) \\ & + s^5(5.805 + 15.1338g_1 + 5.290g_2 + 3.693g_3 + 47.7395g_4 - 0.0003g_5 - 0.0040g_7) \\ & + s^4(3.4528 + 14.5125g_1 + 9.1725g_2 + 6.2054g_3 + 128.3280g_4 + 0.0054g_5 - 0.0003g_6 - 0.0447g_7 \\ & \quad + .5033g_8) \\ & + s^3(6.0887 + 7.3738g_1 + 3.8188g_2 + 5.0478g_3 + 213.4985g_4 + 0.0805g_5 + 0.0054g_6 - 0.0853g_7 \\ & \quad + 5.3707g_8) \\ & + s^2(8.8337 + 1.795g_1 + .2813g_2 + 1.5123g_3 + 169.3315g_4 + .1045g_5 + 0.0805g_6 - 0.0578g_7 + 8.7607g_8) \\ & + s(13.4938 + 0.1825g_1 + 0.080g_2 + 0.1377g_3 + 44.4725g_4 + 0.0328g_5 + 0.1045g_6 - 0.0129g_7 \\ & \quad + 13.4818g_8) \\ & + k(8.4805 + 0.030g_1 + 0.0290g_3 + 0.0328g_6 + 8.4805g_8) \end{aligned} \quad (17)$$

However, we desire that the coefficients of this eighth-order polynomial assume values specified by the optimized denominator D' . Setting the coefficients of the given orders of s in this system description equal to the coefficients of the given orders of s obtained in the optimal form denominator, we generate a set of simultaneous equations in the eight unknowns g_1 to g_8 .

The solution of this set of simultaneous linear equations specifies those feedback gains to be employed in the "actual" system which will implement the previously determined equalization requirements. In this case, the necessary gains are:

$$g_1 = 3.7758$$

$$g_2 = 1.2375$$

$$g_3 = 14.8846$$

$$g_4 = 0.2586$$

$$g_5 = 313.0944$$

$$g_6 = 241.8274$$

$$g_7 = -457.9764$$

$$g_8 = -0.9992$$

This implementation synthesis does indeed produce the selected form of system equalization in terms of the system structure. However, the g_2 loop in this sample case is figurative, and proposes a pick-off of the engine function f_T . An additional figurative feedback loop would be in evidence if a transform for the throttle control had been included in the system structure. Within the context of the present methodology, these figurative loops portray a need for sensors to obtain the indicated function or else networks to generate the indicated function from existing output quantities. These conditions, however, are construed here as problems to be discussed at the engineering level of a given application.

EXPERIMENTAL VALIDATION

The utility of a given system treatment is not measured solely by its analytical bases or by the ease with which it can be implemented. A more rigorous criterion of effectiveness is the enhancement accruing to system performance in an operational environment. Technical as well as financial factors, however, prohibit the conduct of field experiments to fully develop and evaluate system treatments. A less conclusive but tenable index of a treatment's merit can be obtained through the laboratory simulation and experimental comparison of the man-machine system treatments of interest. In the present case, the controlled experimental exercise of several control loop configurations attempts to measure not only the superiority of a given system treatment but also—by implication—the merit of the equalization methodology.

Four experimental conditions or control loop configurations were compared. The first configuration was that of the originally described loop containing a vertical rate meter having a 2-second time constant and labeled the "lag" or \dot{h}_m condition. The second configuration introduced the previously derived equalization of this loop and labeled the "equalized lag" or \dot{h}_{mq} condition. Another possible—but not current—configuration is that of a loop containing an ideal sensor which has no lag time in the sensing and display of vertical rate. This was the third experimental configuration and labeled the "no lag" or \dot{h}_o condition. The fourth experimental configuration was the equalization of this loop to a binomial standard and labeled the "equalized no lag" or \dot{h}_{oq} condition. Figures 11(a) and 11(b) present the analog computer simulation of these four experimental control loop configurations.

Subjects were given a compensatory tracking task in all four control loop configurations. The control was a light, position-type lever, and the display was a 1/16-inch tracking dot presented on a 5-inch oscilloscope. The instructions were to maintain the tracking dot at the center reference of the scope which represented a given vertical velocity. A forcing function composed of two sinusoids (fig. 12) perturbed the loop continuously during test runs of approximately 1 minute. Two experienced trackers served as subjects. After training, each subject received 30 randomly distributed test trials per condition. Test data were first taken on the \dot{h}_m , \dot{h}_o , and \dot{h}_{mq} conditions. At each sitting three trials per condition were given, making up individual sessions of nine trials each. The equalization networks were then modified to create the equalized "no-lag" condition \dot{h}_{oq} . Five sessions, with 10 trials per session,

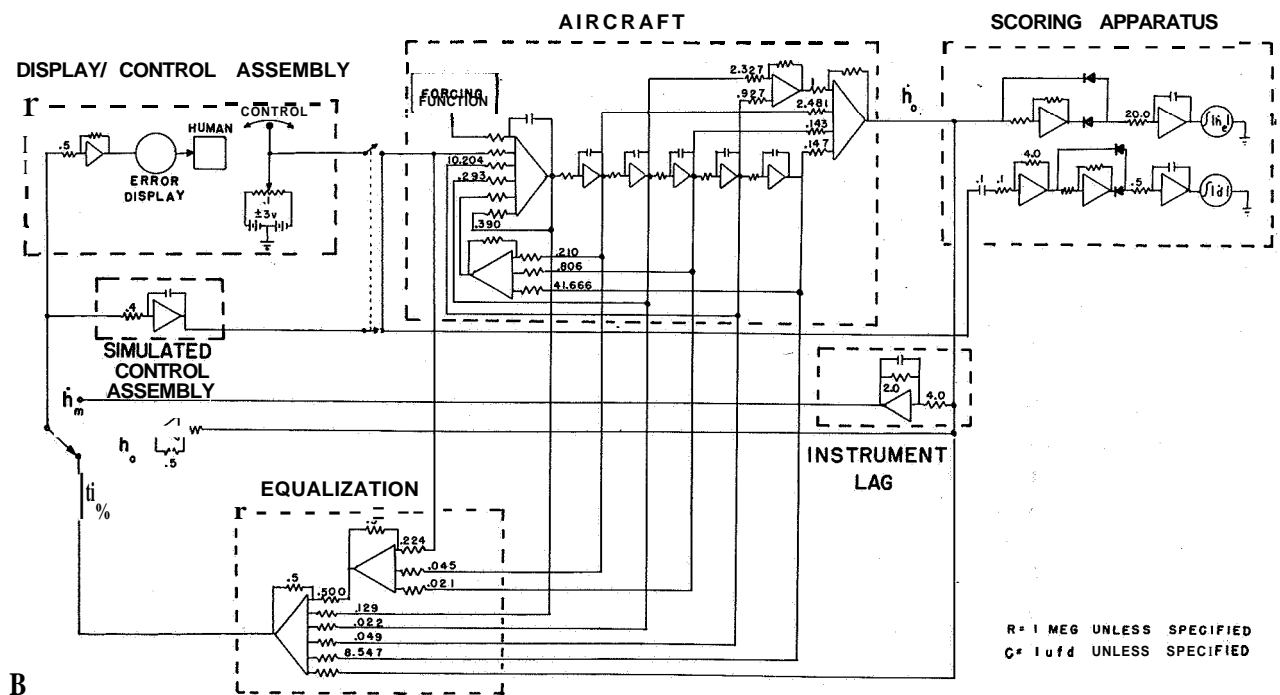
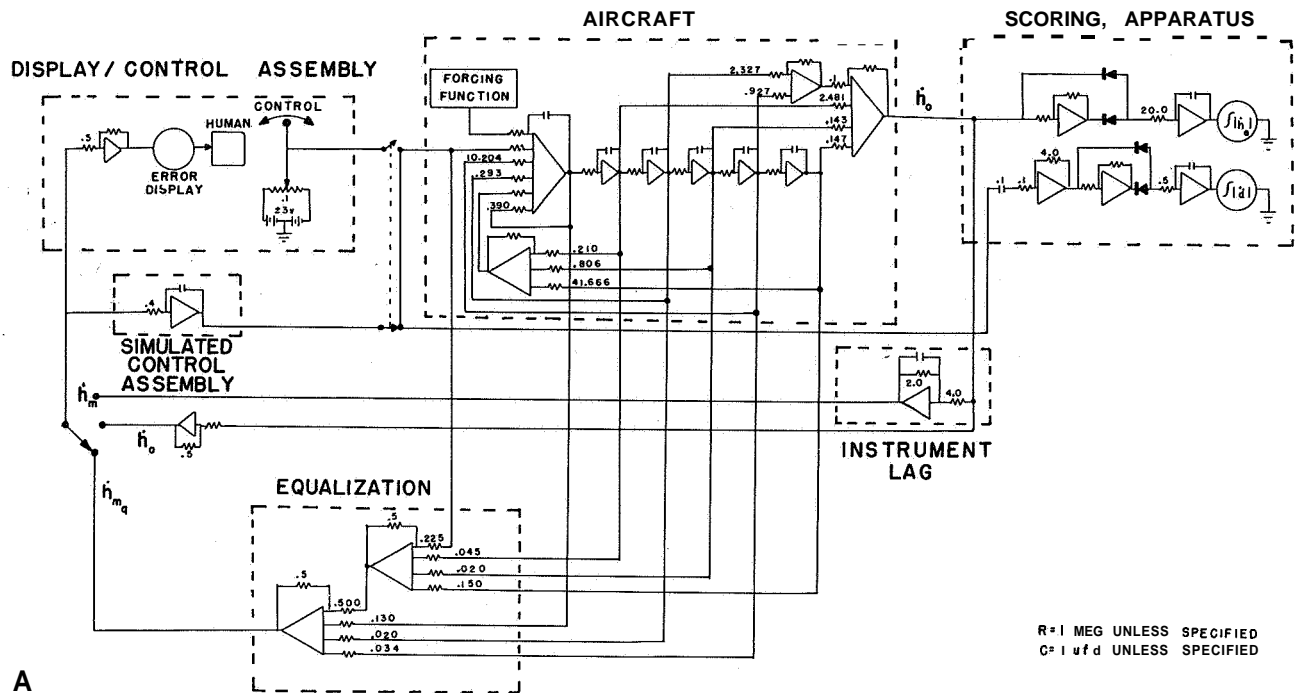


Figure 11.—Analog computer simulator. (a) "Lag" h_m , "no-lag" h_o , and "equalized lag" h_{mq} vertical-rate control-loop configurations; (b) "Lag" h_m , "no-lag" h_o , and "equalized no-lag" h_{mq} vertical-rate control-loop configurations.

were then given to generate 30 test trials on the \dot{h}_{oq} condition. Six trials on \dot{h}_{oq} and two trials each on h_m and h_o , respectively, made up a single session of 10 trials. The ten additional

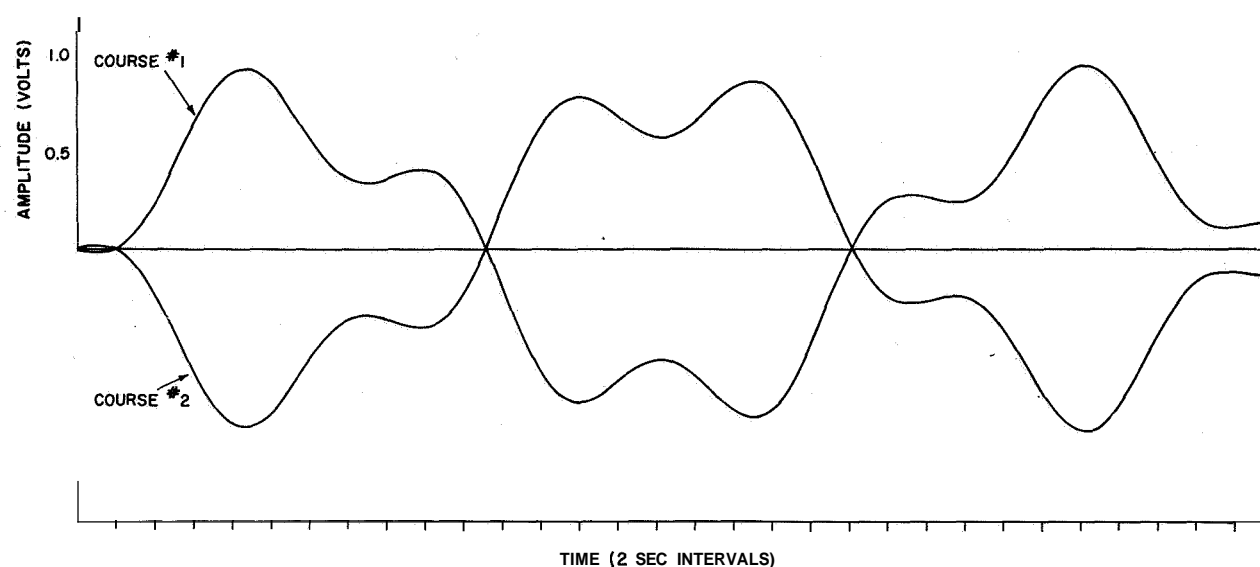


Figure 12.—Sinusoidal forcing functions.

trials cumulated on the h_m and h_o conditions served as check data to insure that equipment and subjects remained comparable under the modified conditions.

The primary measure of system performance was the absolute value of integrated error incurred during each test run. **Error** was defined as any deviation of the system's actual vertical velocity h_o from a fixed value—in this case a value of zero volts. Due to the influence of throttle motion on fuel consumption, as well as on system accuracy, it was deemed desirable to measure the amount of stick manipulation incurred in the different control loops. Thus, a second performance measure was the integrated absolute value of stick rate δ_t .

The results of the experiment are shown in the bargraph of figure 13 and in the statistics of table 1. One way of characterizing these results is to compare the performance scores with those obtained in the given or conventional control loop h_m . First, it will be noted that the employment of a perfect vertical velocity sensor results in a loop configuration h_o which reduces system error by 80 percent but increases stick manipulation by 200 percent. Possibly, this result would be seriously modified in the case of a pilot performing the restrained throttle manipulations which are characteristic of normal flight procedures. Thus, an expected consequence would be much less throttle movement with a concomitant increase in vertical velocity error.

The most important result, in the context of this paper, is the system performance evidenced in the equalized control loops. The k_{mq} and k_q system configurations—both equalized to the binomial standard—gave comparable performance scores resulting in a 90-percent reduction

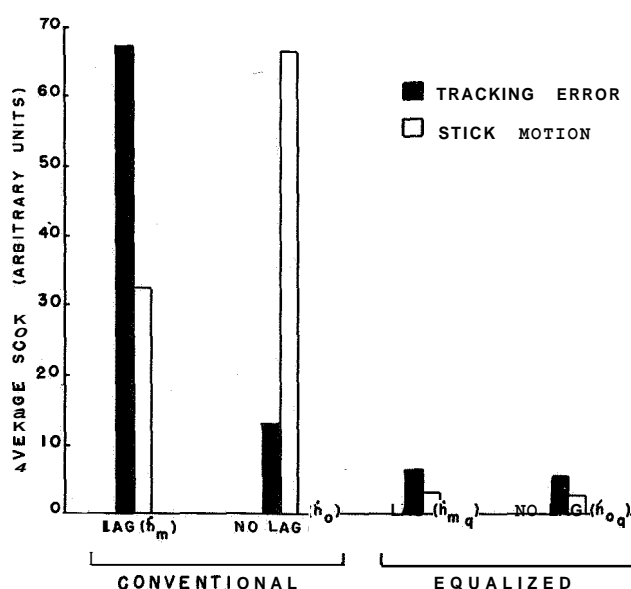


Figure 13.—Average error scores and amounts of stick motion incurred in four vertical velocity control-loop configurations.

TABLE 1.—ANALYSIS OF VARIANCE AND MULTIPLE RANGE TEST
OF MEANS FOR TRACKING ERROR AND STICK MOTION OCCURRING IN
FOUR CONTROL LOOP CONFIGURATIONS

(a) Analysis of variance—tracking error

| Source | df | F |
|-----------|-----|---------------------|
| Loops (L) | 3 | 779.08 ^a |
| Ss (S) | 1 | 4.44 ^b |
| L·S | 3 | 2.61 |
| Within | 232 | |
| Total | 239 | |

^a
p<0.01

^b
p<0.05

Duncan's multiple range test (p<0.01)
(tracking error)

| | | | |
|-------------|-------------|----------------|----------------|
| \dot{h}_m | \dot{h}_o | \dot{h}_{mq} | \dot{h}_{oq} |
| 66.65 | 12.78 | <u>6.41</u> | <u>5.00</u> |

(b) Analysis of variance—stick motion

| Source | df | F |
|-----------|-----|---------------------|
| Loops (L) | 3 | 972.09 ^a |
| Ss (S) | 1 | 38.49 ^a |
| L·S | 3 | 2.44 |
| Within | 232 | |
| Total | 239 | |

^a
p<0.01

Duncan's multiple range test (p<0.01)
(stick motion)

| | | | |
|-------------|-------------|----------------|----------------|
| \dot{h}_o | \dot{h}_m | \dot{h}_{mq} | \dot{h}_{oq} |
| 66.13 | 32.03 | <u>2.71</u> | <u>2.64</u> |

in system error as well as a 90-percent reduction in the amount of stick manipulation. Finally, the results tend to confirm that the presented methodology provides a simple and effective form of system synthesis applicable to the study and design of a very large population of man-machine systems.

REFERENCES

1. Birmingham, H. P.; and Taylor, F. V.: A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems. NRL rep. 4333, Naval Research Laboratory, 1954.
2. Truxal, John G.: Automatic Feedback Control System Synthesis. McGraw-Hill Book Co., Inc., 1965.
3. Thaler, G. J.; and Brown, R. G.: Analysis and Design of Feedback Control Systems. McGraw-Hill Book Co., Inc., 1960.
4. Grabbe, E. M.; et al.: Handbook of Automation, Computation, and Control. Vol. 1, John Wiley & Sons, Inc., 1950.
5. Perry, B. L.; and Birmingham, H. P.: A Method of Synthesizing Higher Order Functions, NRL rep. 5823, Naval Research Laboratory, 1962.
6. Graham, Dunstan; and Lathrop, R. C.: The Synthesis of "Optimum" Transient Response; Criteria and Standard Forms. AIEE, vol. 72, part 11, 1953, pp. 281-282.
7. Wolkovitch, Julian; et al.: Performance Criteria for Linear Constant-Coefficient Systems with Deterministic Inputs. Tech. rep. no. ASD-TR-61-501, Air Force Systems Command, Feb. 1962.

34. Applications of the Pilot Transition Response Model to Flight Control System Failure Analysis

David H. Weir
Systems Technology, Inc.

N68-15935

Manual control situations where the effective dynamics of the task are time varying, and hence the dynamics of the operator are time varying also, are becoming increasingly important in modern aerospace vehicles. There have been a number of past studies involving time-varying dynamics with both slow and rapid changes. This paper concentrates on those with a sudden or instantaneous change between stationary control conditions, with emphasis on the application of a pilot transition model previously derived (ref. 1).

There are a number of practical ways in which the effective dynamics of the task can change suddenly. One class of examples is given by the failure of an artificial stability-augmentation device, such as a pitch or yaw rate damper on an aircraft. Another example is given by perceptual transitions, where the visual field from which the operator derives his control cues changes suddenly in content and form. This occurs, for example, at breakout in an ILS approach when the pilot goes from a heads-down instrument scan to an outside-the-cockpit real-world display of the runway. A similar type of perceptual transition may occur in driving a car, when the driver's gaze is shifted from a preview-like point of regard well ahead of the car, where the effective dynamics are K/s , to a point close alongside the vehicle, where the dynamics are more like K/s^2 .

Current efforts in this area at Systems Technology, Inc., are being directed toward the general problem of making these transitions easier or more "graceful"; that is, minimizing the operator's control difficulties, resultant vehicle motions, time to reach a steady-state adaptation to the new dynamics, and so forth. In essence, operator control techniques and system design criteria which will permit what Licklider once called "graceful degradation" are being sought,

These notions have been expressed in a "graceful degradation hypothesis," which states that the operator will have the least transition control difficulty and his performance will be best if the difference in the effective controlled elements is minimized before and after transition. For example, a transition in task dynamics from a $K \rightarrow K/s^2$ should be worse than from $K/s \rightarrow K/s^2$. One implication of the hypothesis is that the pilot should be given an airplane with less than optimum handling qualities if a failure is likely, and, conversely, a less reliable system can be tolerated if the pilot is required to exert a small amount of control all the time in the unreliable axis.

Part of the current study efforts is directed toward experimental testing of this hypothesis in a fixed-base simulator. The first series of experiments involved a single-axis roll

*The research reported in this paper was sponsored by the Ames Research Center, NASA, under Contract NAS 2-3607.

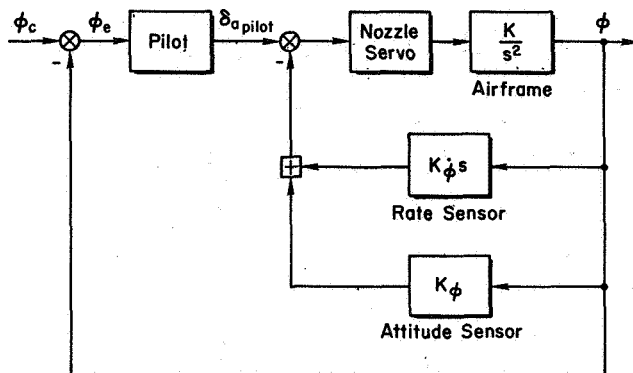


Figure 1.—Block diagram for single-axis roll control task.

backs went suddenly to their maximum value. The frequency of occurrence of a given failure was determined by its unreliability. The failures occurred at random times.

Typical data for $K/s \rightarrow K/s^2$ with a soft mode of failure are shown in figure 2. The failure occurred when the rate feedback went to zero. The error shows one large peak, and the operator's output shows one large postfailure bang.

This suggests that he was nearly time optimal, and corresponds to the response predicted by the current transition model. Some data for $K \rightarrow K/s^2$ show time-optimal characteristics as well, although the magnitude of the first error peak and the divergence rate are larger.

There is a theoretical basis in the current transition model for K/s having smaller error peaks and faster transition times than K/s^2 . This is illustrated in figure 3. Phase plane trajectories for K/s and K/s^2 are shown on the left for time-optimal control following the retention phase, and corresponding time histories are shown on the right. The retention phase ends at time t_1 when the operator has detected the failure according to some criterion (e.g., ref. 2) and is ready to react. He bangs the manipulator c over to some maximum deflection. For $Y_{c_2} = K/s$ this

action causes the error rate to change sign instantaneously, and the operator returns the manipulator to null after the error has been reduced to within a threshold value. For $Y_{c_2} = K/s^2$ the error continues to increase for a short time after t_1 , and then the manipulator has to be reversed at time t_2 when the optimum switching line is

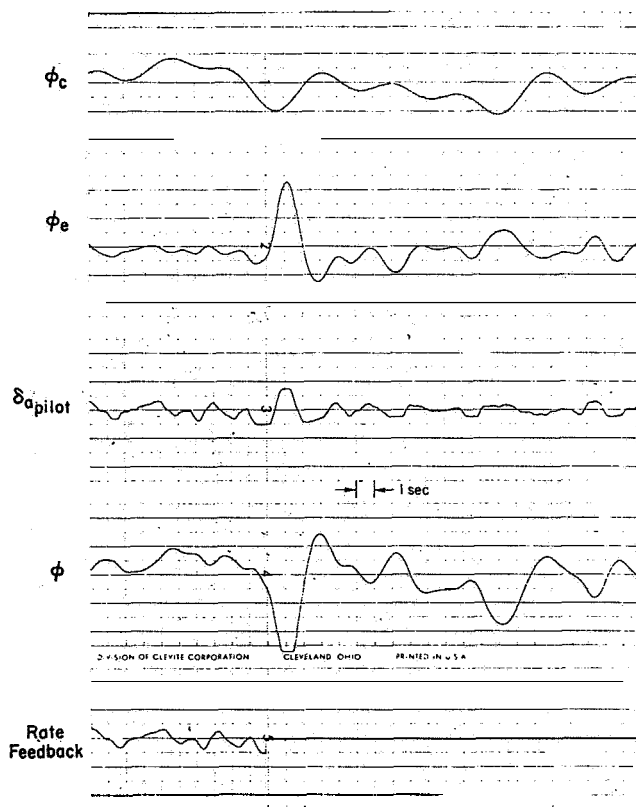


Figure 2.—Transition response data ($\frac{K}{s} \rightarrow \frac{K}{s^2}$, soft).

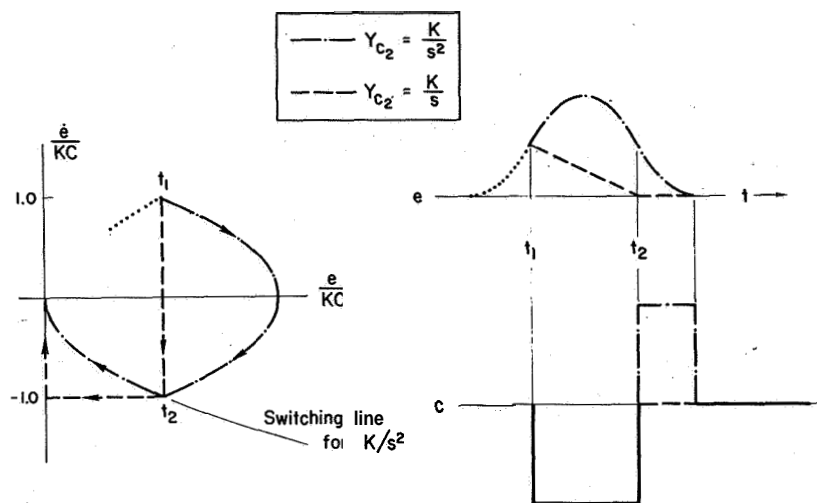


Figure 3.—Comparison of peak errors and transition time.

reached. Following this second bang, he shuts off when the origin is reached. It is clear that the time optimal control assumed in the current transition model gives smaller error peaks and shorter transition times for K/s than for K/s^2 , thereby supporting the graceful degradation hypothesis. A similar result would occur with a quasi-linear form of transition response model.

A second set of experiments is currently in progress involving lateral-directional control of roll angle with aileron and yaw rate with rudder. Various levels of yaw-rate damping are used, and at the time of damper failure the task dynamics become that of a relatively bad airplane.

In summary, all data available at this time support the graceful degradation hypothesis. This, together with the previously derived transition response model, will ultimately allow us to specify rational design criteria which may place more load on the pilot in some flight conditions, thereby relieving some of the requirements for augmentation system redundancy.

REFERENCES

1. Weir, D. H.; and Phatak, A. V.: Model of the Human Operator Response to Step Transitions in Controlled Element Dynamics. NASA CR-671, 1967.
2. Miller, Duncan C.: A Model for the Adaptive Response of the Human Controller to Sudden Changes in Controlled Process Dynamics. B. S. and M. S. Thesis, M. I. T., 1965.